(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 June 2005 (02.06.2005)

PCT

(10) International Publication Number
WO 2005/049642 A2

(51) International Patent Classification[7]: C07K 14/195

(21) International Application Number:
PCT/IB2004/003578

(22) International Filing Date:
23 September 2004 (23.09.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03/13687 21 November 2003 (21.11.2003) FR

(71) Applicants *(for all designated States except US)*: **INSTITUT PASTEUR** [FR/FR]; 25-28, rue du Docteur Roux, F-75015 Paris (FR). **INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE (INSERM)** [FR/FR]; 101, rue de Tolbiac, F-75013 Paris (FR). **UNIVERSITE CLAUDE BERNARD LYON 1** [FR/FR]; 43, boulevard du 11 novembre 1918, F-69100 Villeurbanne (FR). **CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)** [FR/FR]; 3, rue Michel-Ange, F-75016 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BUCHRIESER, Carmen** [AT/FR]; 11 rue de l'Amiral Mouchez, F-75013 Paris (FR). **TICHIT, Magalie** [FR/FR]; Hoff 1, 25 rue de Vouille, F-75015 Paris (FR). **ETIENNE, Jérôme** [FR/FR]; 2, rue Jamen-Grand, F69300 Calvire (FR). **MA, Laurence** [FR/FR]; 29, rue Danton, F-92240 Malakoff (FR). **CAZALET, Christel** [FR/FR]; 27, passage Barrault, F-75013 Paris (FR). **GLASER, Philippe** [FR/FR]; 72 rue de la Glacière, F-75013 Paris (FR). **RUSNIOK, Christophe** [FR/FR]; 44 rue des Gallardons, Batiment A, F-92350 Le Plessis Robinson (FR). **BOUCHIER, Christiane** [FR/FR]; 16, rue Voltaire, F-75011 Paris (FR). **ZIDANE, Nora** [FR/FR]; 31 bis, rue de la Chapelle, F-91310 Montlhery (FR). **MAGNIER, Arnaud** [FR/FR]; 47, avenue du Président Wilson, F-93100 Montreuil (FR). **KUNST, Frédérik** [FR/FR]; 46, rue Barbès, Batiment 18, F-94200 Ivry sur Seine (FR). **VANDENESCH, François** [FR/FR]; 10 rue Berrod, F-69004 Lyon (FR). **JARRAUD, Sophie** [FR/FR]; C/O Université de Lyon, Inserm E 0230 (FR).

(74) Agent: **MARTIN, Jean-Jacques**; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: GENOME OF *LEGIONELLA PNEUMOPHILA* PARIS AND LENS STRAIN-DIAGNOSTIC AND EPIDEMIOLOGICAL APPLICATIONS

(57) Abstract: The object of the invention is the genomic sequence and nucleotidic sequences coding for polypeptides of *Legionella pneumophila* Paris strain and Lens strain, such as cellular surface polypeptides, especially specific between these two strains and/or relative to the Philadelphia strain, or implied in the virulence or in the polysaccharide biosynthesis of cellular envelope, as well as vectors including said sequences and cells transformed by these vectors. The invention also concerns processes for detection of these nucleic acids or polypeptides and diagnostic typing kits for bacteria of the *Legionella* genre, especially of the*Legionella pneumophila* species, such as the Paris and Lens strains, between them and/or relative to the Philadelphia strain. The invention especially concerns a repeated nucleic sequence specific to the *Legionella pneumophila* species and its utilization as an analysis target in processes for detection of the presence of these bacteria. The aim of the invention is also a method for selection of compounds capable of modulating the biosynthesis of these polysaccharides of cellular envelope utilizing said nucleotidic sequences or said polypeptides. The invention finally comprises pharmaceutical compositions, especially vaccinal, for the prevention and/or treatment of bacterial infections, in particular by *Legionella pneumophila* Paris strain and/or Lens strain.

WO 2005/049642 A2

GENOME OF *LEGIONELLA PNEUMOPHILA* PARIS AND LENS STRAIN - DIAGNOSTIC AND EPIDEMIOLOGICAL APPLICATIONS

The object of the invention is the genomic sequence and nucleotidic sequences coding for polypeptides of *Legionella pneumophila* Paris and Lens strain, such as polypeptides of cellular surface, especially specific between these two strains and/or relative to the Philadelphia strain, or implied in the virulence or in the biosynthesis of polysaccharides having a cellular envelope, as well as vectors including said sequences and cells transformed by these vectors. The invention likewise concerns processes for detection of these nucleic acids or polypeptides and diagnostic kits or kits for typing bacteria of the *Legionella* genre, especially the *Legionella pneumophila* species, such as the Paris and Lens strain, between them and/or relative to the Philadelphia strain. The invention especially concerns a specific repeated nucleic sequence of the *Legionella pneumophila* species and its utilization as an analysis target in processes for detection of the presence of these bacteria. The aim of the invention also is a method for selection of compounds capable of modulating the biosynthesis of these polysaccharides having a cellular envelope utilizing said nucleotidic sequences or said polypeptides. The invention finally comprises des pharmaceutical compositions, especially vaccinal, for the prevention and/or treatment of bacterial infections, in particular by the *Legionella pneumophila* Paris and/or Lens strain.

Legionella is a bacteria of the environment responsible for legionellosis and Pontiac fever. The epidemiological data indicate that probably only certain isolates are capable of causing clinical cases. The *L. pneumophila* species seems to have a more significant virulence than the other species, by being responsible for 90% of the cases of legionellosis. At the centre of this species, among the fifteen serogroups, the isolates of serogroup 1 are associated with 80% of cases.

To date, transmission from person to person has never been observed and measures for preventing legionellosis are thus concentrated on elimination of this pathogen from water circulation or from water-cooling towers in air-conditioning systems. In order to establish a rational policy for prevention, it is necessary to prevent the risk associated with each strain. In this optic, it would be desirable to be able to have processes or diagnostic kits available based on the recognition of protein or specific nucleic acid of this genre legio species pathogen or of a particular strain (or again sub-species) of this species.

In effect, the interest in using these specific sequences in the domain of diagnostics or epidemiology rests on the possibility of analyzing a large number of sequences at the same time and very rapidly for:

- classification or typing of bacteria as a function of the presence of a sequence or of a profile of sequences characteristic of a genre, species or strain (sub-species) of bacteria, in particular in association with the gravity or not of pathologies which such bacteria can induce in case of infection in mammals, especially in humans; and

- simultaneous comparison of sequence or profile of sequences between different genres, species or strain (sub-species) of bacteria, pathogenic or not, especially enabling identification of a gene, or the corresponding proteic sequence, or a profile of genes whereof the presence and/or expression in a bacteria is specific to a genre, species or strain of bacteria, and/or to its pathogenicity or not, especially by means of tools such as DNA chips or, if required, protein chips, on which these specific sequences are immobilized.

These specific sequences can be specific sequences of the *Legionella* genre, or of a pathogenic bacteria of the *Legionella* genre and/or of the *Legionella pneumophila* species, or again of a bacteria of the *Legionella pneumophila* species Paris and/or Lens strain or again specific to a bacteria of the *Legionella pneumophila* species Paris and/or Lens strain relative to the Philadelphia strain.

This information is widely utilized especially to rapidly identify the presence or not of the pathogenic bacteria, the gravity of the infection which it can cause, the treatment adapted to an infection, and/or the necessity and the means to be put in place to decontaminate the objects, circuits or fluids which are contaminated or could be contaminated This information will likewise be widely used for epidemiological studies relative to this genre of bacteria.

This is just one of the aspects of the present invention.

In a first stage, the inventors have studied and attempted to comprehend the genomic diversity at the centre of the *Legionella* genre by complete sequencing of the *L. pneumophila* serogroup 1 strain, Paris strain and Lens strain found in different French départements (2) and low-level sequencing of covering of two strains not belonging to the *L. pneumophila* species. The strains selected are *L. longbeachae*, responsible for cases of legionellosis essentially in Australia, and *L. anisa* frequently found in water circulation but not found in patients.

The complete sequencing has likewise been carried out within the scope of the invention of an epidemic *L.pneumophila* serogroup 1 strain known as « Lens strain », responsible for a major epidemic in France with 86 cases and 17 deaths between November 2003 and January 2004.

The genes found to be variable between the strains of *Legionella,* as well as preserved genes having a functional tie to the virulence of *L.pneumophila* can be utilized to manufacture DNA chips.

A large number of isolates of various origins isolated from the environment or originating from clinical cases could be analyzed by using this tool to identify markers enabling the two categories of strains to be discriminated. The comparison of endemic and epidemic isolates will likewise provide bases for understanding the specificities of these strains and in particular the adaptability and stability of the Paris strain.

This approach also helps identify new functions necessary to the virulence of *Legionella* in humans and aid understanding of the different stages of this disease.

The object of the invention is to allow development of novel tools for typing the strains of *Legionella*. These tools could be of the DNA "chip" type or of another type. The novel characteristics of these typing tools will be the following:

* Rapidity and simplicity of use
* High capacity for discrimination between the strains
* Possibility for providing information on the genomic content of the strain analyzed and possibly prevent the risk associated with contamination by *Legionella*.

The inventors have, during this study, brought to light genes found to be variable between the strains of *Legionella*, as well as preserved genes having a functional tie to wall or cellular envelope, or the virulence of *L. pneumophila*, genes which will be able to be used for carrying out these processes or diagnostic kits, especially for producing biochips with protein or DNA.

A large number of isolates of diverse origin isolated from the environment or originating from clinical cases could be analyzed by using these tools for the purpose of identifying markers for discriminating the two categories of strains.

This approach will also help identify new functions necessary for the virulence of *Legionella* in humans and aid in comprehension of the different stages of this disease.

The object of the invention is thus to allow development of novel tools for typing of strains of *Legionella*. These tools could especially be of the protein or

DNA/RNA biochip type. The novel characteristics of these typing tools will be the following:

- rapidity and simplicity of use;
- high capacity for discrimination between strains; and
- possibility of supplying information on the genomic content of the strain analyzed and of possibly preventing the risk associated with contamination by *Legionella*.

The inventors have, in the first instance, sequenced the complete genome of *L. pneumophila* Paris strain in the form of 56 contigs (SEQ ID No. 1 to 56), a sequence made up of a long chromosome of around 3.65 Mb and a long plasmid of around 36 kb. The inventors have likewise identified on these contigs (SEQ ID No. 1 to 56) the nucleic sequences coding for proteins with their respective function (cf Table I with the annotated sequences). The inventors have additionally compared these sequences to the sequences of the genome of *L. pneumophila* Philadelphia strain available at the website http://genome3.cpmc.columbia.edu/~legion/index.html and revealed their presence or not in the sequences of this Philadelphia strain. The sequences of the genome of *L. pneumophila* Philadelphia strain available on the website http://genome3.cpmc.columbia.edu/~legion/index.html correspond to the sequences of the 51 contigs identified in the list of sequences under the SEQ ID Nos. 3456 to 3506. This comparison, made from the available genomic sequence of *L. pneumophila* Philadelphia strain and the proteic sequences obtained by the inventors from the 6 possible reading frames, has revealed that some 88 % of these two genomes are very strongly preserved (95 to 100 % of proteic identity), the remaining 12 % being specific to each strain (cf. Tables I and IV). These results thus demonstrate that there is a wide genomic diversity within the *L. pneumophila* species. A serine protease autotransporter homolog in which is inserted ten repetitions in tandem of a pattern of 60 amino acids was especially identified among the genes specific to the Paris strain. The autotransporters are secretion systems for the Gram negative bacteria in which the N-terminal and C-terminal parts respectively enable secretion across the internal membrane and the formation of pores in the external membrane. The central part of the protein can then remain exposed at the level of the cellular surface or can be split and salted out in the external medium. The role of certain autotransporters in the virulence of the negative Gram bacteria has already been shown; furthermore, work on the autotransporters of the enterobacteria has helped identify a group of serine protease

present solely in pathogenic bacterias whereof the diversity of functions could be linked to the adaptation to the niche occupied by the pathogen.

In addition, the inventors have revealed a very wide inter-species diversity from the sequencing of a pathogenic strain of *L. longbeachae*, which causes very few cases of legionellosis in France, but which is the major source of legionellosis in Australia, and that of a non-pathogenicic *L. anisa* strain. We were able to identify to date 703 ORFs of the *L. longbeachae* strain, whereof 53 % are specific to it; the majority of ORFs preserved between the *L. longbeachae* and *L. pneumophila* Paris strains have a percentage of proteic homology greater than 80%. The preliminary results obtained by the inventors on the non-pathogenic *L. anisa* strain have helped identify 54 % of specific sequences and un percentage of proteic homology greater than 70 % for the sequences preserved between the *L. anisa* and *L. pneumophila* Paris strains.

Tables I and II (« bestblast » obtained for each of the nucleic and proteic sequences corresponding to the annotated ORFs) and X to XXI hereinbelow comprise for each of the ORFs identified either in the Paris strain (Tables I and XIV), or in the Lens strain (Table XVI), its position on the contigs or chromosomes, and, if required, the existence of a peptide signal, the best result of the blast on nrprot (Best-Blastp). The ORFs:

- specific to the *L. pneumophila* Paris strain relative to the *L. pneumophila* Philadelphia strain;
- specific to the *L. pneumophila* Paris strain relative to the *L. pneumophila* Philadelphia and Lens strains (Table XVII);
- specific to the *L. pneumophila* Lens strain relative to the *L. pneumophila* Philadelphia and Paris strains (Table XVIII);
- specific to the *L. pneumophila* Philadelphia strain relative to the *L. pneumophila* Paris and Lens strains (Table XI);

were identified in considering as specific the ORFs having a percentage of proteic identity less than 75 %. In the event where ORF is preserved in the two genomes, the percentage of identity between the two proteins is mentioned.

In this Table I, the ORFs present in the partial sequence of the *L. longbeachae* strain have likewise been noted. Finally, the ORFs specific to the *Legionella* genre were identified by considering as specific the ORFs having a percentage of proteic identity with sequences of the nrprot bank less than 25 %.

In Table XIX, the ORFs present at the same time in the Paris and Lens strain were indicated, though absent in the Philadelphia strain.

In Table XX, the ORFs present at the same time in the Paris and Philadelphia strain were indicated, though absent in the Lens strain.

Finally, in Table XXI, the ORFs present at the same time in the Philadelphia and Lens strain were indicated, though absent in the Paris strain.

In conclusion, the diversity revealed by the content of the present patent application helps define proteic probes, such as antibodies, or DNA probes for developing a typing tool. The utilization of this tool on a large number of strains isolated from patients and strains isolated from the environment will enable this tool to be validated, a tool which will aid in predicting the risk associated with a strain by discriminating in a certain manner the strains isolated from patients of other strains.

Among the significant families of proteins of *Legionella pneumophila* Paris strain the family of surface proteins or that of proteins implied in the biosynthesis of surface polysaccharides can be cited, or again that of proteins implied in the virulence of these bacteria. The process of evolution has allowed the development of a number of unique mechanisms on the Gram+bacteria, by which they can immobilize proteins on their surface. The functions of these different proteins of cellular walls are extremely diverse. However, many proteins linked covalently to the surface of the Gram+ pathogens are estimated to be important for the survival of the pathogen inside the infected host. The study of *Legionella pneumophila* Paris strain demands novel approaches, in particular genetic, to improve understanding of the different metabolic paths of this organism.

Accordingly, it is object of the present invention to divulge the complete sequence of the genome of *Legionella pneumophila* Paris strain (Collection de the Institut Pasteur CIP 107-629-T), a sequence obtained from a collection of clones (BAC) filed on 19 November 2003 with the Collection Nationale de Cultures de Microorganisms (CNCM) [National Collection of Microorganism Cultures], 25 rue du Docteur Roux, 75724 Paris Cedex 15, France, according to the arrangements of the Budapest Treaty and registered under file number I-3137, as well as all the genes contained in said genome.

It is also another object of the present invention to divulge the complete sequence of the genome of *Legionella pneumophila* Lens strain, a sequence obtained from a collection of clones (BAC) filed on 23 September 2004 with the Collection

Nationale de Cultures de Microorganisms (CNCM) [National Collection of Microorganism Cultures], 25 rue du Docteur Roux, 75724 Paris Cedex 15, France, according to the arrangements of the Budapest Treaty and registered under file number I-3306, as well as all the genes contained in said genome.

In effect, knowledge of the genome of these organisms enables the interactions between the different genes, the different proteins, and the different metabolic paths to be better defined. In effect, and contrary to divulging isolated sequences, the complete genomic sequence of an organism forms a whole, allowing all the information necessary to this organism to grow and function to be obtained immediately.

If the present invention provides the nucleotidic sequence of the genome of *Legionella pneumophila* Paris strain (Collection de the Institut Pasteur CIP 107-629-T), having been the object of a filing of a collection of clones (BAC) covering this genome with the C.N.C.M. in Paris on 19 November 2003 and registered under file number I-3138, and likewise provides the nucleotidic sequence of the genome of *Legionella pneumophila* Lens strain, a sequence obtained from a collection of clones (BAC) filed on 23 September 2004 with CNCM, and likewise provides certain polypeptide sequences coded by these two genomes, the specialist will be able to determine the other ORFs, by utilizing known methods, and appropriate software.

In the set of claims hereinbelow, the term « nucleotidic sequence » will especially be able to be replaced by the term « polynucleotide » without modifying the object and the scope of the set of claims such as filed.

The present invention thus relates to:

- a genomic nucleotidic sequence of *Legionella pneumophila* Paris strain characterized in that it is selected among the sequences SEQ ID 3507 and 3508, SEQ ID N° 55 and the sequences SEQ ID N° 1 to SEQ ID N° 54, and SEQ ID N° 56;

- a genomic nucleotidic sequence of *Legionella pneumophila* Lens strain characterized in that it is selected among the sequences SEQ ID 6733 and 6734.

The present invention likewise relates to an isolated or purified nucleotidic sequence:

(A) of *Legionella pneumophila* Paris strain, characterized in that it is selected among:

a) a nucleotidic sequence comprising at least one sequence having 80 % identity with the sequences SEQ ID 3507 and 3508, and SEQ ID N° 1 to SEQ ID N° 56;

b) a nucleotidic sequence hybridizing in very stringent conditions with the sequences SEQ ID 3507 and 3508, and SEQ ID N° 1 to SEQ ID N° 56;

c) a nucleotidic sequence complementing sequences SEQ ID 3507 and 3508 and SEQ ID N° 1 to SEQ ID N° 56, or complementing a nucleotidic sequence such as defined in a), or b), or a corresponding nucleotidic sequence of RNA; and

d) a nucleotidic sequence of at least 15 nucleotides of fragment representative of sequences SEQ ID 3507 and 3508, and SEQ ID N° 1 to SEQ ID N° 56, or of fragment representative of their sequence, or

(B) of *Legionella pneumophila* Lens strain, characterized in that it is selected among:

a) a nucleotidic sequence comprising at least one sequence having 80 % of identity with the sequences SEQ ID 6733 and 6734;

b) a nucleotidic sequence hybridizing in very stringent conditions with the sequences SEQ ID 6733 and 6734;

c) a nucleotidic sequence complementing sequences SEQ ID 6733 and 6734; or complementing a nucleotidic sequence such as defined in a), or b), or a corresponding nucleotidic sequence of RNA; and

d) a nucleotidic sequence of at least 15 nucleotides of fragment representative of sequences SEQ ID 6733 and 6734; or of fragment representative of their sequence.

More particularly, the object of the present invention likewise is the nucleotidic sequences characterized in that they originate from sequences SEQ ID 3507 and 3508, and SEQ ID N° 1 to SEQ ID N° 56, and in that they code for a polypeptide selected from amongst the polypeptides of sequences SEQ ID 3509 to SEQ ID 6732, and SEQ ID N° 56 to SEQ ID N° 3455, preferably coding for a secreted enzyme likewise present in the Lens strain and Philadelphia strain of sequences SEQ ID 3675, 4267, 4292 and 6477, preferably coding for a polypeptide present on the surface of *Legionella pneumophila* Paris strain of sequence SEQ ID Nos. 3410, 704, 746, 2267, 2751, 3192, 3218, 3221, 3222, 3317, 3324, 136, 171, 310, 337, 481, 527 652, 664, 893, 972, 1148, 1298, 1361, 1503, 1521, 1576, 1651, 1755, 1847, 1877, 2224, 2406, 2843, 2930, 3037, 3139, 3157, 3165, 3181, preferably coding for a polypeptide present on the specific surface of *Legionella pneumophila* Paris strain relative to the Philadelphia strain, especially of sequence SEQ ID Nos. 3410, 171, 337, 481, 652, 1148, 1521, 2843, 3037, 3181, or one of its representative fragments of at least 5 amino acids, or coding for a

polypeptide implied in the biosynthesis of polyssacharides having a cellular envelope of sequence SEQ ID Nos. 1126, 3218, 288, 632, 917, 1503, 1555, 1877, 1928, 1963, 2204, 2212, 2243, 2324, 2378, 2410, 2411, or again the nucleic sequences of the rtxA gene, especially those coding for the polypeptides of sequences SEQ ID Nos. 3410, 3037, 3165 and 3181.

The present invention also relates more generally to the nucleotidic sequences derived from the sequences SEQ ID Nos. 3507 and 3508, and SEQ ID N° 1 to SEQ ID N° 56, and coding for a polypeptide of *Legionella pneumophila* Paris strain such that they can be isolated from these sequences SEQ ID Nos. 3507 and 3508, and SEQ ID N° 1 to SEQ ID N° 56.

In addition, the nucleotidic sequences characterized in that they comprise a nucleotidic sequence selected among:

a) a nucleotidic sequence coding for a polypeptide selected from amongst the sequences SEQ ID Nos. 3509 to 6732, and SEQ ID N° 56 to SEQ ID N° 3455;

b) a nucleotidic sequence comprising at least 80 %, 85 %, 90 %, 95 % or 98 % of identity with a nucleotidic sequence coding for a polypeptide selected from amongst the sequences SEQ ID Nos. 3509 to 6732, and SEQ ID N° 56 to SEQ ID N° 3455;

c) a nucleotidic sequence hybridizing in very stringent conditions with a nucleotidic sequence coding for a polypeptide selected from amongst the sequences SEQ ID Nos. 3509 to 6732, and SEQ ID N° 56 to SEQ ID N° 3455;

d) a complementary nucleotidic sequence or RNA corresponding to a sequence such as defined in a), b) or c);

e) a nucleotidic sequence of a representative fragment having at least 15 nucleotides of a sequence such as defined in a) or d); and

f) a modified nucleotidic sequence of a sequence such as defined in a), d) or e),

are likewise objects of the invention.

Nucleic acid, nucleic sequence or nucleic acid, polynucleotide, oligonucleotide, polynucleotidic sequence, nucleotidic sequence, terms which will be employed indifferently in the present description, are understood to designate precise chaining of nucleotides, modified or not, effectively defining a fragment or a region of a nucleic acid, comprising or not non-natural nucleotides, and able to correspond just as well to a double-strand DNA, a single-strand DNA as transcription products of said DNAs. Therefore, the nucleic sequences according to the invention likewise encompass the PNA (Peptid Nucleic Acid), or similar.

It must be understood that the present invention does not relate to the nucleotidic sequences in their natural chromosomic environment, that is, in the natural state. These are sequences which were isolated and/or purified, that is they were sampled directly or indirectly, for example by copy, their environment having been at least partially modified. It is understood to likewise designate the nucleic acids obtained by chemical synthesis.

« Percentage of identity » between two sequences of nucleic acids or amino acids in the sense of the present invention is understood to designate a percentage of nucleotides or residues of identical amino acids between the two sequences to be compared, obtained after the best alignment, this percentage being purely statistical and the differences between the two sequences being distributed randomly and over their entire length. "Best alignment" or " optimal alignment " is understood to designate the alignment for which the percentage of identity determined hereinafter is the highest. The comparisons of sequences between two sequences of nucleic acids or amino acids are traditionally made by comparing these sequences after they were aligned in optimal fashion, said comparison being made by segment or by « window of comparison » to identify and compare the local regions of similarity of sequence. The optimal alignment of the sequences for comparison can be made, apart from manually, by means of the local of de Smith and Waterman (1981, Ad. App. Math. 2:482), by means of the local homology algorithm of Neddleman and Wunsch (1970, J. Mol. Biol. 48:443), by means of the similarity search method of Pearson and Lipman (1988, Proc. Natl. Acad. Sci. USA 85:2444), by means of software utilizing these algorithms (GAP, BESTFIT, BLAST P, BLAST N, FASTA and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI). To obtain optimal alignment, the program BLAST is preferably used, with the BLOSUM 62 matrix. The PAM or PAM250 matrices can likewise be used.

The percentage of identity between two sequences of nucleic acids or amino acids is determined by comparing these two sequences aligned optimally in which the sequence of nucleic acids or amino acids to be compared can comprise additions or deletions relative to the reference sequence for optimal alignment between these two sequences. The percentage of identity is calculated by determining the number of identical positions for which the nucleotide or the residue of amino acid is identical between the two sequences, by dividing this number of identical positions by the total

number of positions compared and by multiplying the result obtained by 100 to obtain the percentage of identity between these two sequences.

Nucleic sequences having a percentage of identity of at least 80 %, preferably 85 % or 90 %, more preferably 95 % or 98 % or 99 %, after optimal alignment with a reference sequence are understood to designate he nucleic sequences exhibiting, relative to the nucleic reference sequence, certain modifications such as in particular deletion, truncation, elongation, chimeric fusion and/or substitution, especially specific, and whereof the nucleic sequence has at least 80 %, preferably 85 %, 90 %, 95 %, 98 % or 99 %, of identity after optimal alignment with the nucleic reference sequence. These are preferably sequences whereof the complementary sequences are capable of being hybridized specifically with the reference sequences. Preferably, the specific hybridization conditions or stringent conditions will be such that they ensure at least 80 %, preferably 85 %, 90 %, 95 %, 98 % or 99 % of identity after optimal alignment between one of the two sequences and the sequence complementary to the other.

Hybridization in very stringent conditions signifies that the temperature and ionic force conditions are selected in such a way that they permit hybridization to be maintained between two fragments of complementary DNA. By way of illustration, very stringent conditions of the hybridization stage for the purpose of defining the polynucleotide fragments described hereinabove, are advantageously the following.

DNA-DNA or DNA-RNA hybridization is performed in two stages: (1) prehybridization at 42°C for 3 hours in phosphate buffer (20 mM, pH 7.5) containing 5 x SSC (1 x SSC corresponds to a solution 0.15 M NaCl + 0.015 M sodium citrate), 50 % formamide, 7 % sodium dodecyl sulfate (SDS), 10 x Denhardt's, 5 % dextran sulfate and 1 % DNA salmon sperm; (2) hybridization per se for 20 hours at a temperature depending on the size of the probe (i.e.: 42°C, for a probe of size > 100 nucleotides) followed by 2 washes of 20 minutes at 20°C in 2 x SSC + 2 % SDS, 1 wash of 20 minutes at 20°C in 0.1 x SSC + 0.1 % SDS. The last wash is done in 0.1 x SSC + 0.1 % SDS for 30 minutes at 60°C for a probe of size > 100 nucleotides. The very stringent hybridization conditions described hereinabove for a polynucleotide of defined size can be adapted by the specialist for oligonucleotides of larger or smaller size, according to the teaching of Sambrook *et al.*, (1989, Molecular cloning: a laboratory manual. 2nd Ed. Cold Spring Harbor).

In addition, fragment representative of sequences according to the invention is understood to designate any nucleotide fragment having at least 15 nucleotides,

preferably at least 20, 25, 30, 50, 75, 100, 150, 300 and 450 consecutive nucleotides of the sequence from which it originates.

Representative fragment is understood in particular to be a nucleic sequence coding for a biologically active fragment of a polypeptide, such as defined hereinbelow.

Representative fragment is likewise understood to be the intergenic sequences, and in particular the nucleotidic sequences bearing the regulation signals (promoters, terminators, or enhancers, ...), or again probe or primer sequences aiding in specifically detecting or amplifying the nucleic sequences coding for the polypeptides of sequences SEQ ID Nos. 3509 to 6732, and SEQ ID N° 56 to SEQ ID N° 3455.

Of said representative fragments those are preferred having nucleotidic sequences corresponding to open reading frames, known as ORF sequences (ORF for « Open Reading Frame »), included in general between an initiation codon and a stop codon, or between two stop codons, and coding for polypeptides, preferably of at least 100 amino acids, such as for example, without limiting them, the ORF sequences to be described hereinafter.

The representative fragments according to the invention can be obtained for example by specific amplification such as PCR or after digestion by appropriate restriction enzymes of nucleotidic sequences according to the invention, this method being described in particular in the work of Sambrook *et al.*. Said representative fragments can likewise be obtained by chemical synthesis as long as their size is not too significant, according to methods well known to the specialist.

The representative genome fragments of *Legionella pneumophila* Paris strain according to the invention likewise comprise at least one fragment of at least 15 nucleotides or more as cited above for the fragments resulting from enzymatic cutting at the level of a restriction site. Of course, expression proteins such as RNA or proteins are understood according to the present invention.

Among the sequences containing inventive sequences, or representative fragments, are likewise understood the sequences which are naturally framed by sequences which present at least 80 %, 85 %, 90 %, 95 % or 98 % of identity with the sequences according to the invention.

Modified nucleotidic sequence is understood as any nucleotidic sequence obtained by mutagenesis according to techniques well known to the specialist, and comprising des, preferably a maximum 10 %, 7.5 %, 5 %, 2.5 %, 1 %, 0.5 %, 0.1 % or even less than 0.01 %, of modified nucleotides, relative to normal sequences, for

example mutations in the regulating and/or promoting sequences of the expression of the polypeptide, especially leading to modification of the rate of expression or activity of said polypeptide.

Modified nucleotidic sequence is likewise understood as any nucleotidic sequence coding for a polypeptide modified such as defined hereinbelow.

The representative fragments according to the invention can likewise be probes or primers, which can be utilized in processes for detection, identification, dosage or amplification of nucleic sequences.

In a preferred manner the invention is relative to a nucleotidic sequence coding for a polypeptide according to the invention.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a specific polypeptide of a bacteria of the *Legionella* genre, or one of its fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a specific polypeptide of a pathogenic bacteria of the *Legionella* genre and/or of the *Legionella pneumophila* species, or one of its fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a specific polypeptide of a bacteria of the *Legionella pneumophila* species Paris strain, or one of its fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a specific polypeptide of a bacteria of the *Legionella pneumophila* species Paris strain relative to the Philadelphia strain, or one of its fragments of at least 5 amino acids, in particular selected from amongst the polypeptides of sequence SEQ ID Nos. 3410, 171, 337, 481, 652, 1148, 1521, 2843, 3037, 3181 or one of its fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a specific polypeptide of a bacteria of the *Legionella pneumophila* species Paris strain relative to the Lens and Philadelphia strains, or one of its fragments of at least 5 amino acids, in particular

selected from amongst the polypeptides whereof the sequences are indicated in Table XVII or one of their fragments of at least 5 amino acids, or their complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a surface polypeptide of *Legionella pneumophila* Paris strain, or one of its fragments of at least 5 amino acids, in particular selected from amongst the sequence polypeptides SEQ ID Nos. 3410, 704, 746, 2267, 2751, 3192, 3218, 3221, 3222, 3317, 3324, 136, 171, 310, 337, 481, 527 652, 664, 893, 972, 1148, 1298, 1361, 1503, 1521, 1576, 1651, 1755, 1847, 1877, 2224, 2406, 2843, 2930, 3037, 3139, 3157, 3165, 3181, or one of its fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a polypeptide of specific surface of *Legionella pneumophila* Paris strain relative to the Philadelphia strain, selected from amongst the polypeptides of sequences SEQ ID Nos. 3410, 171, 337, 481, 652, 1148, 1521, 2843, 3037, 3181, or one of its fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a polypeptide implied in the biosynthesis of polysaccharide having a cellular envelope of *Legionella pneumophila* Paris strain, in particular selected from amongst the polypeptides of sequence SEQ ID Nos. 1126, 3218, 288, 632, 917, 1503, 1555, 1877, 1928, 1963, 2204, 2212, 2243, 2324, 2378, 2410, 2411, or one of its representative fragments of at least 5 amino acids, or its complementary nucleic sequence.

In a preferred manner the invention is relative to a nucleotidic sequence according to the invention, characterized in that it codes for a polypeptide of *Legionella pneumophila* Paris strain coded by the rtxA gene of sequence SEQ ID Nos. 3410, 3037, 3165 and 3181, or one of its representative fragments of at least 5 amino acids, or its complementary nucleic sequence.

The rtxA gene of Legionella pneumophila was demonstrated as being implied in virulence. This gene codes for a protein of 1208 aa. Four ORFs whereof the references are given hereinbelow (SEQ ID 3410 – 3037 – 3165 – 3181) correspond to this gene in the Paris strain which would code for a protein of at least 4000 aa. Comparison with the Philadelphia strain shows the presence of a homologous gene for the N- and C-terminal

parts. The central part of the protein is constituted by repetitions of a pattern of around 20 aa different between the two strains.

In another aspect, in a preferred manner, the invention is relative to a nucleotidic sequence according to the invention characterized in that it codes for a polypeptide of *Legionella pneumophila* Paris strain, or one of its representative fragments of at least 5 amino acids implied in the biosynthesis of the amino acids, in the biosynthesis of the cofactors, prosthetic and transporter groups, implied in the cellular machinery, implied in the central intermediate metabolism, implied in the energetic metabolism, implied in the metabolism of fatty acids and phospholipids, implied in the metabolism of nucleotides, purines, pyrimidines or nucleosides, implied in the functions of regulation, implied in the process of replication, implied in the process of transcription, implied in the process of translation, implied in the process of transport and bonding of proteins, implied in adaptation to atypical conditions, implied in sensitivity to medications and the like, or implied in the functions relative to transposons.

Owing to the genomic sequence presented in the present invention, the specialist will know how to identify the genes coding for proteins regulating transcription of the genes of *Legionella pneumophila* Paris strain. In addition, Table I provides the list of the open reading phases (ORF for « Open Reading Frame) annotated and identified on the genome of *Legionella pneumophila* Paris strain (SEQ ID N° 1 to SEQ ID N° 55), with especially their position on said genome, and, in Table II, the putative functions which can be attributed to them by utilizing customary techniques for comparing the genomic (« Bestblast »). All the same, such a list must not be considered as limiting, where one protein can have several roles in the cell.

Modifying the structure or the integrity of these genes could help modify the expression of the target genes controlled by the target promoters of these regulators. Thus, the expert will be able to select the regulator(s) pertinent for the required application as well as their target, thus allowing optimization of the expression of genes of interest. The utilization of the tools described above such as the DNA chips, also registers all the genes whereof the regulation is modified by inactivation of certain genes. It is thus possible to select a set of control sequences responding to the same type of regulation. These sequences can then be used to control the expression of genes of interest.

In general, the list of sequences SEQ ID, or their corresponding coding nucleic sequence could be determined by the specialist from the most probable putative

functions determined for each of the sequences SEQ ID in Tables I and XIV hereinbelow for each of the classes of activity classified hereinbelow.

It is important to note all the same that a living organism is a whole and must be taken as such. Accordingly, so as to develop and exhibit its properties, any organism has need of interaction between the different metabolic paths. Therefore, the abovementioned classification must not be considered as limiting, with a gene able to be implied in two distinct metabolic paths.

The invention likewise relates to polypeptides coded by a nucleotidic sequence according to the invention, preferably by a fragment representative of the sequences SEQ ID Nos. 3509 to 6732, the sequence SEQ ID N° 55 or sequences SEQ ID N° 1 to SEQ ID N° 54, and SEQ ID N° 56, and corresponding to an ORF sequence, such as described in Table XIV (coding for one of the sequences SEQ ID N° 3509 to SEQ ID N° 6732), and in Table I (coding for one of the sequences SEQ ID N° 57 to SEQ ID N° 3455).

In particular, polypeptides of *Legionella pneumophila* Paris strain, characterized in that they are selected from amongst the following polypeptides:

- polypeptides of sequences SEQ ID Nos. 3509 to 6732 and of sequences SEQ ID N° 56 to SEQ ID N° 3455;
- preferably enzymes secreted by *Legionella pneumophila* Paris, Lens and Philadelphia strains, especially sequences SEQ ID Nos. 3675, 4267, 4292 and 6477;
- preferably polypeptides present on the surface of *Legionella pneumophila* Paris strain, especially of sequence SEQ ID Nos. 3410, 704, 746, 2267, 2751, 3192, 3218, 3221, 3222, 3317, 3324, 136, 171, 310, 337, 481, 527 652, 664, 893, 972, 1148, 1298, 1361, 1503, 1521, 1576, 1651, 1755, 1847, 1877, 2224, 2406, 2843, 2930, 3037, 3139, 3157, 3165, 3181, and still more preferred polypeptides present on the specific surface of *Legionella pneumophila* Paris strain relative to the Philadelphia strain, especially those of sequences SEQ ID Nos. 3410, 171, 337, 481, 652, 1148, 1521, 2843, 3037, 3181, or one of its representative fragments of at least 5 amino acids;
- polypeptides implied in the biosynthesis of polysaccharides having a cellular envelope, especially of sequence SEQ ID Nos. 1126, 3218, 288, 632, 917, 1503, 1555, 1877, 1928, 1963, 2204, 2212, 2243, 2324, 2378, 2410, 2411;
- or again polypeptides of sequence SEQ ID Nos. 3410, 3037, 3165 and 3181, coded by the rtxA gene.

The invention likewise comprises polypeptides characterized in that they comprise a polypeptide selected from amongst:

a) a polypeptide of sequences SEQ ID Nos. 3509 to 6732, and of sequences SEQ ID N° 56 to SEQ ID N° 3455;

b) a polypeptide having at least 80 % preferably 85 %, 90 %, 95 % and 98 % of identity with an polypeptide of sequences SEQ ID Nos. 3509 to 6732, and of sequences SEQ ID N° 56 to SEQ ID N° 3455;

c) a fragment of at least 5 amino acids, preferably biologically active, of one such as defined in b);

d) a biologically active fragment of a polypeptide of sequence SEQ ID N° 56 to SEQ ID N° 3455; and

e) a modified polypeptide of a polypeptide of sequences SEQ ID Nos. 3509 to 6732, and of sequences SEQ ID N° 56 to SEQ ID N° 3455, comprising at most 10 % modified amino acids, preferably 7.5 %, 5 %, 2.5 %, 1 %, 0.5 %, 0, 1 % or again 0.01 %.

The nucleotidic sequences coding for the abovedescribed polypeptides are likewise an object of the invention.

In the present description, the terms polypeptides, polypeptide sequences, peptides and proteins are interchangeable.

It must be understood that the invention does not relate to polypeptides in the natural form, that is, that they are not taken in their natural environment, rather they were isolated or obtained by purification from natural sources, or else obtained by genetic recombination, or by chemical synthesis, and that they can then comprise non-natural amino acids such as will be described hereinbelow.

Polypeptide having a certain percentage of identity with another, likewise designated by homologous polypeptide, is understood to designate those polypeptides having, relative to natural polypeptides, certain modifications, in particular deletion, addition or substitution of at least one amino acid, truncation, elongation, a chimeric solution and/or a mutation, or polypeptides having post-translational modifications. Of the homologous polypeptides preference is given to those whereof the sequence of amino acids has at least 80 %, preferably 85 %, 90 %, 95 %, 98 % and 99 % of identity with the sequences of amino acids of the polypeptides according to the invention. In the case of substitution, one or more consecutive or non-consecutive amino acid(s) are replaced by « equivalent » amino acids. The expression « equivalent amino acids» in

this case endeavors to designate any amino acid likely to be substituted for one of the amino acids of the base structure without however essentially modifying the biological activities of the corresponding peptides and such as they will be defined hereinbelow.

These equivalent amino acids can be determined either by being supported on their structural homology with the amino acids for which they are substituted, or on results of comparative assays of biological activity between the different polypeptides capable of being effected.

By way of example, mention is made of the possibilities of substitution for being carried out without the resulting extensive modification of the biological activity of the corresponding modified polypeptide. Thus leucine can be replaced by valine or isoleucine, aspartic acid by glutamine acid, glutamine by asparagine, arginine by lysine, etc., the inverse substitutions naturally being envisaged under the same conditions.

The homologous polypeptides correspond likewise to the polypeptides coded by the homologous or identical nucleotidic sequences, such as defined previously and thus comprise in the present definition mute polypeptides or corresponding to inter or intra species variations, able to exist in *Legionella*, and which correspond especially to truncations, substitutions, deletions and/or additions, of at least one residue of amino acids.

It is understood that the percentage of identity between two polypeptides is calculated in the same way as between two sequences of nucleic acids. Thus, the percentage of identity between two polypeptides is calculated after optimal alignment of these two sequences, on a maximum homology window. To define said maximum homology window, the same algorithms as for the nucleic acid sequences can be utilized.

Biologically active fragment of a polypeptide according to the invention is understood to mean in particular a fragment of polypeptide, such as defined hereinbelow, having at least one of the biological characteristics of the polypeptides according to the invention, especially in that it capable of exerting in general an even partial activity, such as for example:

- enzymatic (metabolic) activity or an activity able to be implied in the biosynthesis or biodegradation of organic or inorganic compounds;
- structural activity (cellular envelope, coping molecule, ribosome);
- transport activity (energy, ion); or in the secretion of protein;

- activity in the process of replication, amplification, preparation, transcription, translation or maturation, especially of DNA, RNA or proteins.

Fragment of polypeptides according to the invention is understood to mean a polypeptide comprising a minimum of 5 amino acids, preferably 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 75, 100, 150 acids and 200 amino acids.

The fragments of polypeptides can correspond to isolated or purified fragments naturally present in the strains of *Legionella*, or to fragments which can be obtained by cleavage of said polypeptide by a proteolitic enzyme such as trypsine or chymotrypsine or collagenase, by a chemical reagent (cyanogen bromide, CNBr) or by placing said polypeptide in a highly acidic environment (for example at pH = 2.5). Polypeptidic fragments can likewise be prepared by chemical synthesis, from hosts transformed by a vector of expression according to the invention which contains a nucleic acid allowing expression of said fragment, and placed under the control of the appropriate elements of regulation and/or expression.

« Modified polypeptide» of a polypeptide according to the invention is understood to mean a polypeptide obtained by genetic recombination or by chemical synthesis such as described below, which has at least one modification relative to the normal sequence, preferably at most 10 % of amino acids modified relative to the normal sequence, preferably even at most 7.5 %, 5 %, 2.5 %, 1 %, 0.5 %, 0, 1 % or again 0.01 %. These modifications can be especially made to amino acids necessary for the specificity or efficacy of the activity, or at the origin of the structural conformation, the charge, or the hydrophobicity of the polypeptide according to the invention. Polypeptides of equivalent, augmented or diminished activity, or of equivalent, narrower or wider specificity can thus be created. Among the polypeptides modified, those polypeptides in which up to five amino acids can be modified, truncated at the N end or C terminal, or else deleted, or added, must be mentioned.

As is indicated, the object of the modifications of a polypeptide especially are:

- to permit its usage in biosynthesis or biodegradation processes of organic or inorganic compounds,
- to permit its usage in processes of replication, amplification, repair and transcription, translation, or maturation especially of DNA, RNA, or proteins,
- to permit its improved secretion,
- to modify its solubility, efficacy or specificity of activity, or again to facilitate its purification.

Chemical synthesis likewise has the advantage of being able to utilize non-natural amino acids or non-peptidic bonds. Therefore, it can be interesting to utilize non-natural amino acids, for example in the D form, or analogs of amino acids, especially sulphurized forms.

In another aspect, the invention is preferably relative to a polypeptide according to the invention, characterized in that it is a specific polypeptide of a bacteria of the *Legionella* genre, or one of its fragments of at least 5 amino acids.

In another aspect, the invention is preferably relative to a polypeptide according to the invention, characterized in that it is a specific polypeptide of a pathogenic bacteria of the *Legionella* genre and/or of the *Legionella pneumophila* species, or one of its representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the present invention, characterized in that it is a specific polypeptide of a bacteria of the species *Legionella pneumophila* Paris strain, or one of its representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the present invention, characterized in that it is a specific polypeptide of a bacteria of the *Legionella pneumophila* species Paris strain relative to the Philadelphia strain, or one of its representative fragments of at least 5 amino acids, in particular selected from amongst the polypeptides of sequences SEQ ID Nos. 3410, 171, 337, 481, 652, 1148, 1521, 2843, 3037, 3181, or one of its representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the present invention, characterized in that it is a specific polypeptide of a bacteria of the *Legionella pneumophila* species Paris strain relative to the Lens and Philadelphia strains, or one of its representative fragments of at least 5 amino acids, in particular selected from amongst the polypeptides whereof the sequence is indicated in Table XVII, or one of their representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the present invention, characterized in that it is a surface polypeptide of *Legionella pneumophila* Paris strain, or one of its representative fragments of at least 5 amino acids, in particular selected from amongst the polypeptides of sequence SEQ ID Nos. 3410, 704, 746, 2267, 2751, 3192, 3218, 3221, 3222, 3317, 3324, 136, 171, 310, 337, 481, 527 652, 664, 893, 972, 1148, 1298, 1361, 1503, 1521, 1576, 1651, 1755, 1847,

1877, 2224, 2406, 2843, 2930, 3037, 3139, 3157, 3165, 3181, or one of its representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the present invention, characterized in that it is a polypeptide of specific surface of *Legionella pneumophila* Paris strain relative to the Philadelphia strain, selected from amongst the polypeptides of sequence SEQ ID Nos. 3410, 171, 337, 481, 652, 1148, 1521, 2843, 3037, 3181, or one of its representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the invention, characterized in that it is a polypeptide *Legionella pneumophila* Paris strain implied in the biosynthesis of polysaccharide having a cellular envelope of *Legionella pneumophila* Paris strain, in particular selected from amongst the polypeptides of sequence SEQ ID Nos. 1126, 3218, 288, 632, 917, 1503, 1555, 1877, 1928, 1963, 2204, 2212, 2243, 2324, 2378, 2410, 2411, or one of its representative fragments of at least 5 amino acids.

In another aspect, preferably, the invention is relative to a polypeptide according to the invention, characterized in that it is a polypeptide de *Legionella pneumophila* Paris strain coded by the rtxA gene de sequence SEQ ID Nos. 3410, 3037, 3165 and 3181, or one of its representative fragments of at least 5 amino acids, or its complementary nucleic sequence.

In another aspect, preferably, the invention is relative to a polypeptide according to the invention, characterized in that it is a polypeptide of *Legionella pneumophila* Paris strain, or one of its representative fragments of at least 5 amino acids, implied in the biosynthesis of amino acids, in the biosynthesis of cofactors, prosthetic and transporter groups, implied in the cellular machinery, implied in the central intermediary metabolism, implied in the energetic metabolism, implied in the metabolism of fatty acids and phospholipids, implied in the metabolism of nucleotides, purins, pyrimidins or nucleosides, implied in the functions of regulation, implied in the process of replication, implied in the process of transcription, implied in the process of translation, implied in the process of transport and binding of proteins, implied in the adaptation to atypical conditions, implied in the sensitivity to drugs and the like, or implied in the functions relatives to transposons.

The object of the present invention is likewise the nucleotidic sequences and/or polypeptides according to the invention, characterized in that said sequences are registered on a registration support whose form and nature facilitate reading, analysis

and/or exploitation of said sequence(s). These supports can likewise contain other information extracted from the present invention, especially the analogies with already known sequences, and/or information concerning the nucleotidic sequences and/or polypeptides of other microorganisms for the purpose of facilitating comparative analysis and exploitation of the results obtained.

Among said registration supports, preference is given in particular to those supports readable by a computer, such as magnetic, optic, electric or hybrid supports, in particular information discs, CD-ROM, servers. Such registration supports are likewise an object of the invention.

The registration supports according to the invention, with the information contributed, are very useful for the choice of primers or nucleotidic probes for determining genes in *Legionella pneumophila* Paris strain or strains close to this organism. Similarly, the utilization of these supports for the study of genetic polymorphism of a strain close to *Legionella pneumophila* Paris strain, in particular by determination of the colinearity regions, is very useful as far as these supports provide not only the nucleotidic sequence of the genome of *Legionella pneumophila* Paris strain, but likewise the genomic organization in said sequence. Thus, the uses of registration supports according to the invention are likewise objects of the invention.

The homology analysis between different sequences is completed in effect advantageously by means of software for sequence comparisons, such as BlastP or BlastN software, or other software well known to the specialist.

A probe or primer is defined, in the sense of the invention, as being a fragment of single-strand nucleic acids or a denatured double-strand fragment comprising for example 12 bases at several kb, especially 15 at several hundreds of bases, preferably from 15 to 50 or 100 bases, and possess a hybridization specificity in conditions determined to form a hybridization complex with a nucleic acid target.

The probes and primers according to the invention can be marked directly or indirectly by a radioactive or non-radioactive compound using methods well known to the specialist, to obtain a detectable and/or quantifiable signal.

The non-marked sequences of polynucleotides according to the invention can be utilized directly as probe or primer.

The sequences are generally marked to obtain sequences utilizable for numerous applications. The marking of the primers or probes according to the invention is done by radioactive elements or by non-radioactive molecules.

Examples of the radioactive isotopes used are ^{32}P, ^{33}P, ^{35}S, ^{3}H or ^{125}I. The non-radioactive entities are selected among ligands such as biotin, avidin, streptavidin, dioxygenin, haptenes, colorants, luminescent agents such as radioluminescent, chemoluminescent, bioluminescent, fluorescent, phosphorescent agents.

The polynucleotides according to the invention can thus be utilized as primer and/or probe in processes in particular making use of the PCR technique (amplification in chain by polymerase) (Rolfs *et al.*, 1991, Berlin: Springer-Verlag). This technique requires the choice of pairs of oligonucleotide primers framing the fragment to be amplified. For example, reference can be made to the technique described in the U.S. patent N° 4,683,202. The amplified fragments can be identified, for example after electrophoresis in agarose or polyacrylamide gel, or according to a chromatographic technique such as filtration on gel or ion exchange chromatography, then sequenced. The specificity of the amplification can be controlled by using as primer the inventive nucleotidic sequences of polynucleotides as matrix, plasmids containing these sequences or even the derived amplification products. The amplified nucleotide fragments can be utilized as reagents in hybridization reactions so as to reveal the presence, in a biological sample, of a nucleic acid target of sequence complementary to those of said amplified nucleotide fragments.

The aim of the invention is likewise the nucleic acids capable of being obtained by amplification by means of primers according to the invention.

Other amplification techniques for the nucleic acid target can be advantageously employed as an alternative to PCR (PCR-like) by means of a couple of primers of nucleotidic sequences according to the invention. PCR-like is understood to mean all the methods making use of direct or indirect reproductions of the sequences of nucleic acids, or else in which the marking systems were amplified; these techniques are well known, and in general are amplification of DNA by a polymerase; when the original sample is a RNA reverse transcription should be previously carried out. There are currently numerous processes enabling this amplification, such as for example the SDA technique (Strand Displacement Amplification) or brine displacement amplification technique (Walker *et al.*, 1992, Nucleic Acids Res. 20:1691), the TAS technique (Transcription-based Amplification System) described by Kwoh *et al.* (1989, Proc. Natl. Acad. Sci. USA, 86:1173), the 3SR technique (Self-Sustained Sequence Replication) described by Guatelli *et al.* (1990, Proc. Natl. Acad. Sci. USA 87:1874), the NASBA technique (Nucleic Acid Sequence Based Amplification) described by Kievitis *et al.*

(1991, J. Virol. Methods, 35:273), the TMA technique (Transcription Mediated Amplification), the LCR technique (Ligase Chain Reaction) described by Landegren *et al.* (1988, Science 241:1077), the RCR technique (Repair Chain Reaction) described by Segev (1992, Kessler C. Springer Verlag, Berlin, New-York, 197-205), the CPR technique (Cycling Probe Reaction) described by Duck *et al.* (1990, Biotechniques, 9:142), the Q-beta-replicase amplification technique described by Miele *et al.* (1983, J. Mol. Biol., 171:281). Certain of these techniques have since been refined.

In the event where the polynucleotide target to be detected is a RNAm, prior to use of an amplification reaction by means of the primers according to the invention or prior to use of a detection process by means of the inventive probes, an enzyme of inverse transcriptase type is used advantageously in order to obtain a DNAc from the RNAm contained in the biological sample. The DNAc obtained will then serve as target for the primers or the probes used in the process of amplification or detection according to the invention.

The technique of hybridization of probes can be executed in various ways (Matthews *et al.*, 1988, Anal. Biochem., 169:1-25). The most general method consists of immobilizing the nucleic acid extracted from cells of different tissues or cells in culture on a support (such as nitrocellulose, nylon, polystyrene) and of incubating, in well-defined conditions, the nucleic acid target immobilized with the probe. After hybridization, the probe excess is eliminated and the hybrid molecules formed are detected by the appropriate method (measuring of radioactivity, fluorescence or enzymatic activity associated with the probe).

In accordance with another operating mode of the nucleic probes according to the invention, the latter can be utilized as capture probes. In this case, a probe, known as «capture probe», is immobilized on a support and serves to capture via specific hybridization the nucleic acid target obtained from the biological sample to be tested and the nucleic acid target is then detected due to a second probe, known as «detection probe», marked by an easily detectable element.

Of the possibly interesting fragments of nucleic acids, anti-sense oligonucleotides should thus be cited in particular, that is, whereof the structure ensures, via hybridization with the sequence target, inhibition of the expression of the corresponding product. Sense oligonucleotides which, by interaction with proteins implied in regulating the expression of the corresponding product, will cause either inhibition or activation of this expression, should likewise be cited.

In a preferred way the probes or primers according to the invention are immobilized on a support, covalently or non-covalently. In particular, the support can be a DNA chip or a high-density filter, likewise objects of the present invention.

The interest in using DNA chips or, if required, protein chips in the domain of diagnostics or epidemiology rests as mentioned previously on the possibility of analyzing a large number of sequences at the same time and very rapidly, for:

- classification or typing of bacteria as a function of the presence of a sequence or of a profile of sequences characteristic of the genre, especially of the pathogenicity or not of the genre, especially *Legionella*, or of the species, especially *Legionella pneumophila*, or specific to a bacteria of the *Legionella* genre and/or of the *Legionella pneumophila* species, or specific to a bacteria of the *Legionella pneumophila* sub-species Paris strain, or specific to a bacteria of the *Legionella pneumophila* sub-species Paris strain relative to the Philadelphia strain and/or Lens strain, or even specific to a bacteria of the *Legionella pneumophila* sub-species Paris and Lens strain relative to the Philadelphia strain, in particular, in association with the gravity or not of the pathologies which such bacteria can induce in case of infection in mammals, especially in humans; or

- simultaneous comparison of sequence or of profile of sequences between different genres, species or strain of bacteria, pathogenic or not, allowing especially identification of a gene, or the corresponding proteic sequence, or a profile of genes whereof the presence and/or the expression in a bacteria is specific according to its genre, its species or its sub-species or strain of bacteria, and/or its pathogenicity or not. This information is largely useful especially for rapidly identifying the presence or not of a pathogenic bacteria, the gravity of the infection it can cause, the treatment adapted to an infection, and/or the necessity and the means for implementing contaminated circuits or fluids or able to be contaminated for decontaminating the objects. This information will likewise be largely useful to epidemiological studies relative to this genre of bacteria.

DNA chip or high-density filter is understood to mean a support on which DNA sequences are fixed, each of them able to be marked by its geographic location. These chips or filters differ principally in their size, the material of the support, and possibly the number of DNA sequences fixed thereto.

The probes or primers according to the first invention can be fixed on solid supports, in particular the DNA chips, by means of different fabrication processes. In

particular, *in situ* synthesis can be carried out by photochemical addressing or via ink jet. Other techniques consist of carrying out *ex situ* synthesis and fixing the probes on the support of the DNA chip by mechanical or electronic addressing, or by ink jet. These different processes are well known to the specialist.

In effect, numerous techniques or devices for analysis of biological samples were developed in recent years, in particular for parallel analysis of several quantities of nucleic acids, especially following the development of the genomic.

Among these techniques or devices, the supports enabling high-rate analysis of nucleic acids, such as biochips, or DNA chips (also called « micro- or macroarrays », or even « DNA chip ») were the object of numerous studies.

These biochips can be made in particular from a support, generally solid and functionalized, on which given nucleic acids (nucleic probes) were fixed by covalent bond and localized, and on which nucleic probes are fixed specifically respectively by matching (or specific hybridization) or by recognition of an affinity site of the nucleic acids which are to be detected or identified in the biological sample.

Particular examples of the documents describing techniques relative to bioDNA chips are:

- the review article by Wang J. (Nucleic Acids Research, 28, 16:3011-3016, 2000), which has an abstract making the point on the main known techniques relative to DNA chips,

- the patent document issued under N° US 6,030,782, which describes grafting with a mercaptosilanized surface, of nucleic acids modified by a sulhydryl or disulfide group, and the article by Bamdad (Biophysical Journal, 75:1997-2003, 1998), which describes obtaining surfaces having DNA by incorporation of composite molecules, DNA-thiols, in auto-assembled monolayers (« self-assembled monolayers or SAMs »);

- the international patent application published under N° W0 00/43539 which proposes immobilizing molecules, such as oligonucleotides, by means of polyfunctional polymers (« polymer brushes ») thus enabling the grafting density to be increased. These polymers can be obtained from hydroxyethyl, acrylamide methacrylate, or vinyl pyrrolidone;

- the international patent application published under N° WO 00/36145 describes a fabrication method of DNA chips, comprising polymerization on a substrate of metallic layer type, a copolymer of pyrrol and functionalized pyrrol, fixing a reticulation agent on the functionalized pyrrol, then fixing a biological probe (such as an

oligonucleotide). The reticulation agent can be bifunctional, and for example have an ester function of the N-hydroxysuccinimide and a maleimide function;

- the international patent application published under N° WO 98/20020 which likewise describes the high-density immobilization of nucleic acids on solid supports, this time by placing in contact of a nucleic acid containing a thiol group with support having a group reacting with this thiol, possibly by way of a reticulation agent;

- the article by Penchovsky *et al.* (Nucleic Acids Research, 28, 22, e98, 2000), which describes a method for immobilization of oligonucleotides on aminated balls of latex, by means of a reticulation agent which reacts under the action of light; and

- the international patent applications published under N° WO 99/16907, WO 00/40593 and WO 00/44939 filed by the company Surmodics (which produces lames for registering oligonucleotides functionalized with an amine). These applications describe especially the fixing of nucleic acids on surfaces such as glass, by way of a polymer skeleton to which one or more « photochemically active » groups are fixed on one side of the polymer (for grafting on the surface) and « thermochemically active » on the other side (for grafting with the functionalized nucleic acid).

A nucleotidic sequence (probe or primer) according to the invention thus enables detection and/or amplification of specific nucleic sequences. In particular, detection of said sequences is made easier when the probe is fixed on a DNA chip, or to a high-density filter.

The utilization of DNA chips or of high-density filters in effect helps determine expression of genes in an organism having a genomic sequence close to the genome of *Legionella pneumophila* Paris strain (Collection of the Pasteur Institute CIP 107-629-T).

The genomic sequence of *Legionella pneumophila* Paris strain, completed by identification of all the genes of this organism, such as presented in the present invention, serves as a base for the construction of these DNA chips or filter.

The preparation of these filters or chips consists of synthesizing oligonucleotides, corresponding to the 5' and 3' ends of the genes. These oligonucleotides are selected by utilizing the genomic sequence and its annotations divulged by the present invention. The matching temperature of these oligonucleotides at the corresponding places on the DNA must be approximately the same for each oligonucleotide. This aids in preparing fragments of DNA corresponding to each gene by the utilization of condition of appropriate PCR in a highly automated environment.

The amplified fragments are then immobilized on filters or supports made of glass, silicon or synthetic polymers and these media are utilized for hybridization.

The availability of such filters and/or chips and of the annotated corresponding genomic sequence allows study of the expression of large ensembles, or of the totality of genes in the microorganisms of the *Legionella* genre or of the *Legionella pneumophila* species, by preparing the complementary DNA, and by hybridizing it to the DNA or to the oligonucleotides immobilized on the filters or the chips. Likewise, the filters and/or the chips allow study of the variability of the strains or species, by preparing the DNA of these organisms and by hybridizing it to the DNA or to the oligonucleotides immobilized on the filters or the chips.

The differences between the genomic sequences of the different strains or species can extensively affect the intensity of the hybridization and, as a consequence, perturb interpretation of the results. It can thus be necessary to have the precise sequence of the genes of the strain to be studied. The method for detecting the genes described in detail hereinbelow, implying determination of the sequence of random fragments of a genome, and organizing it according to the sequences of the complete genome of the *Legionella pneumophila* strain divulged in the present invention, can be very useful.

The nucleotide, or proteic, sequences according to the invention can be likewise utilized in DNA chips, or, if required, protein chips for performing mutation analysis. This analysis is based on the constitution of chips, especially DNA chips, capable of analyzing each base of a nucleotidic sequence according to the invention. For this purpose the techniques of micro-sequencing on DNA chip especially could be used. The mutations are detected by extension of immobilized primers hybridizing analyzed sequences to the matrix, just in a position adjacent to that of the desired mute nucleotide. A single-strand matrix, RNA or DNA, of the sequences for analysis will be advantageously prepared according to classic methods, from products amplified according to PCR-type techniques. The matrices of single-strand DNA, or RNA thus obtained are then deposited on the DNA chip, in conditions enabling their specific hybridization to the immobilized primers. A thermostable polymerase, for example Tth or the Taq DNA polymerase, specifically extends the end 3' of the primer immobilized with an analog of complementary marked nucleotide of the nucleotide in a variable position of the site; for example a thermal cycle is created in the presence of fluorescent dideoxyribonucleotides. The experimental conditions will be adapted especially to the

chips employed, to the immobilized primers, to the polymerases employed, and to the marking system selected. An advantage of microsequencing, relative to techniques based on hybridization of probes, is that it allows all the variable nucleotides to be identified with optimal discrimination in conditions of homogeneous reactions; when used on DNA chips, it permits optimal resolution and specificity for routine and industrial detection of mutations in multiplex.

The utilization of high-density filters and/or chips thus provides new knowledge on the regulation of genes in organisms of industrial importance, and in particular the legionelloses propagated in diverse conditions. It also allows rapid identification of the differences between the genome of the strains utilized in multiples industrial applications.

In addition, a DNA chip or a filter can be an interesting extremely tool for determination, detection and/or identification of a microorganism. Therefore, the DNA chips according to the invention which further contain at least one nucleotidic sequence of a microorganism other than *Legionella pneumophila* Paris strain, immobilized on the support of said chip are likewise preferred. Preferably, the selected microorganism is selected from among the species of bacteria of the *Legionella* genre (hereinbelow designated at times as bacteria associated with *Legionella pneumophila*), or the sub-species of *Legionella pneumophila*, or again the variants of *Legionella pneumophila* Paris strain.

A DNA chip or a filter according to the invention is a very useful element of certain kits or is necessary for the detection and/or identification of microorganisms, in particular the bacteria belonging to the *Legionella pneumophila* species Paris strain or the associated microorganisms, likewise objects of the invention.

Besides, the DNA chips or the filters according to the invention, containing probes or specific primers of *Legionella pneumophila* Paris strain, compared in particular to the *Philadelphia* strain, are highly advantageous elements of kits or necessary for the detection and/or quantification of the expression of genes of *Legionella pneumophila* Paris strain (or of associated microorganisms).

In effect, control of the expression of genes is a critical point for optimizing the growth and yield of a strain, either allowing the expression of one or more novel genes, or by modifying the expression of genes already present in the cell. The present invention provides all the sequences naturally active in *Legionella pneumophila* Paris strain enabling the expression of the genes. It thus allows all the sequences expressed in

Legionella pneumophila Paris strain to be determined. It likewise provides a tool for locating the genes whereof the expression follows a given pattern. To achieve this, the DNA of all or part of the genes of *Legionella pneumophila* Paris strain can be amplified thanks to primers according to the invention, then fixed to a support, such as for example glass or nylon or a DNA chip, so as to construct a tool allowing the expression profile of these genes to be followed. This tool, constituted by this support containing the coding sequences serves as hybridization matrix to a mixture of marked molecules reflecting the RNA messengers expressed in the cell (in particular the probes marked according to the invention). By repeating this experience at different instants and by combining all these data via suitable processing, the expression profiles of all these genes are obtained. The knowledge of the sequences which follow a given regulation pattern can also be of benefit for direct searching, for example by homology, for other sequences following globally, but slightly different for the same regulation pattern. By way of complement, it is possible to isolate each control sequence present upstream of the segments serving as probes and to follow their activity by means of appropriate means such as a reference gene (luciferase, β-galactosidase, GFP for « Green Fluorescent Protein »). These isolated sequences can then be modified and assembled by metabolic engineering with sequences of interest with a view to their optimal expression.

The aim of the invention likewise is the cloning and/or expression vectors, which contain a nucleotidic sequence according to the invention. In particular, those nucleotidic sequences are preferred which code for polypeptides having a cellular envelope or surface, or implied in the cellular machinery, in particular secretion, central intermediary metabolism, in particular production of sugar, energetic metabolism, the synthesis process of Vitamin B12, transcription and translation, synthesis of polypeptides.

The vectors according to the invention preferably comprise elements which permit the expression and/or secretion of the nucleotidic sequences in a determined host cell.

The vector must comprise a promoter, signals for initiation and termination of translation, thus of the appropriate regulation regions of transcription. It must be able to be kept stable in the host cell and can possibly have particular signals which specify secretion of the translated protein. These different elements are selected and optimized

by the specialist as a function of the cellular host utilized. To this effect, the nucleotidic sequences according to the invention can be inserted into autonomous replication vectors within the selected host, or they can be integrative vectors of the selected host.

Such vectors are prepared by methods currently utilized by the specialist, and the resulting clones can be introduced to an appropriate host using standard methods, such as lipofection, electroporation, thermal choc, or chemical methods.

The vectors according to the invention are for example vectors of plasmidic or viral origin. They are useful for transforming host cells so as to clone or express the nucleotidic sequences according to the invention.

The invention likewise comprises the host cells transformed by a vector according to the invention.

The cellular host can be selected from amongst prokaryotic or eukaryotic systems, for example bacterial cells, but likewise yeast cells or animal cells, in particular the cells of mammals. The cells of insects or plant cells can likewise be used here. The host cells preferred according to the invention are in particular prokaryotic cells, preferably bacteria such as *E. coli,* or again belonging to the *Legionella* genre, to the *Legionella pneumophila* species Paris strain, or microorganisms associated with the *Legionella pneumophila* species Paris strain.

The invention relates likewise to animals, except human, which comprise a cell transformed according to the invention. The transformed cells according to the invention are utilizable in preparation processes for recombinant polypeptides according to the invention. The preparation processes of a polypeptide according to the invention in recombinant form, characterized in that they utilize a vector and/or a cell transformed by a vector according to the invention are themselves included in the present invention. Preferably, a cell transformed by a vector according to the invention is cultivated in conditions permitting expression of said polypeptide and said recombinant peptide is recovered. The host cells according to the invention can likewise be utilizes for preparation of nutritive compositions, themselves an object of the present invention.

As was mentioned, the cellular host can be selected from amongst prokaryotic or eukaryotic systems. In particular, it is possible to identify nucleotidic sequences according to the invention, facilitating secretion in such a prokaryotic or eukaryotic system. A vector according to the invention carrying such a sequence can thus be used advantageously for production of recombinant proteins, to be secreted. In fact,

purification of these recombinant proteins of interest will be facilitated by the fact that they are present in the supernatant of the culture cellular rather than inside the host cells.

The polypeptides according to the invention can likewise be prepared by chemical synthesis. Such a preparation process is likewise an object of the invention. The specialist is aware of synthesis chemical processes, for example techniques utilizing solid phases (see especially Steward *et al.*, 1984, Solid phase peptides synthesis, Pierce Chem. Company, Rockford, 111, 2nd ed., (1984)) or techniques utilizing partial solid phases, by condensing fragments or by synthesis in classic solution. The polypeptides obtained by chemical synthesis and capable of comprising corresponding non-natural amino acids are likewise included in the invention.

The invention is further relative to hybrid polypeptides having at least one polypeptide or one of its fragments according to the invention, and a sequence of a polypeptide for inducing an immune response in humans or animals.

Advantageously, the antigenic determinant is such that it is capable of inducing a humoral and/or cellular response.

Such a determinant could comprise a polypeptide or one of its fragments according to the invention in glycosylated form utilized with a view to obtaining immunogenic compositions capable of inducing synthesis of antibodies directed against multiple epitopes. Said polypeptides or their glycosylated fragments likewise part of the invention.

These hybrid molecules can be made up in part by a carrier molecule of polypeptides or of their fragments according to the invention, associated with a possibly immunogenic part, in particular an epitope of the diphtheric toxin, tetanic toxin, a surface antigen of the hepatitis B virus (patent FR 79 21811), the antigen VP1 of the poliomyelitus virus or any other toxin or viral or bacterial antigen.

The synthesis processes of hybrid molecules encompass the methods utilized in genetics to construct hybrid nucleotidic sequences coding for the desired polypeptidic sequences. For example, reference can be made advantageously to the technique for obtaining genes coding for fusion proteins described by Minton in 1984.

Said hybrid nucleotidic sequences coding for a hybrid polypeptide, as well as the hybrid polypeptides according to the invention characterized in that these are recombinant polypeptides obtained by the expression of said hybrid nucleotidic sequences, are likewise part of the invention.

The invention likewise comprises the vectors characterized in that they contain one of said hybrid nucleotidic sequences. The host cells transformed by said vectors, the transgenic animals comprising one of said transformed cells as well as the preparation processes for recombinant polypeptides utilizing said vectors, said transformed cells and/or said transgenic animals are likewise naturally part of the invention.

The coupling between a polypeptide according to the invention and an immunogenic polypeptide can be made chemically, or biologically. Therefore, according to the invention, it is possible to introduce one or more binding element(s), especially amino acids for facilitating the coupling reactions between the polypeptide according to the invention, and the immunostimulator polypeptide, the covalent coupling of the immunostimulator antigen able to be formed at the N or C-terminal end of the polypeptide according to the invention. The bifunctional reagents allowing this coupling are determined as a function of the end selected for making this coupling, and the coupling techniques are well known to the specialist.

The conjugates originating from a coupling of peptides can likewise be prepared by genetic recombination. The (conjugated) hybrid peptide can in effect be produced by recombinant DNA techniques, by insertion in or addition to the DNA sequence coding for the polypeptide according to the invention, of a sequence coding for the antigenic, immunogenic or haptenic peptide(s). These preparation techniques for hybrid peptides by genetic recombination are well known to the specialist (see for example Makrides, 1996, Microbiological Reviews 60:512-538).

Preferably, said immune polypeptide is selected in the group of peptides containing anatoxins, especially diphteric toxoid or tetanic toxoid, the proteins derived from Streptococcus (as the binding protein to human seralbumin), the membranous OMPA proteins and the complexes of proteins of external membranes, the vesicles of external membranes or the proteins of thermal shocks.

The hybrid polypeptides according to the invention are very useful to obtain monoclonal or polyclonal antibodies, capable of specifically recognizing the polypeptides according to the invention. In effect, a hybrid polypeptide according to the invention allows potentiation of the immune response, against the polypeptide according to the invention coupled to the immunogenic molecule. Such monoclonal or polyclonal antibodies, their fragments, or chimeric antibodies, recognizing the polypeptides according to the invention, are likewise objects of the invention.

The specific monoclonal antibodies can be obtained according to the classic method of hybridome culture described by Köhler and Milstein (1975, Nature 256, 495).

The antibodies according to the invention are for example chimeric antibodies, humanized antibodies, Fab, or $F(ab')^2$ fragments. It can likewise be in the form of immunoconjugate or antibodies marked so as to provide a detectable and/or quantifiable signal.

Therefore, the antibodies according to the invention can be employed in a process for detection and/or identification of bacteria belonging to the *Legionella pneumophila* species Paris and/or Lens strain and/or Philadelphia, or to an associated microorganism in a biological sample, characterized in that it comprises the following stages:

a) contact of the biological sample with an antibody according to the invention;

b) evidence of the possibly formed antigen-antibody complex.

The antibodies according to the present invention are likewise utilizable for detecting an expression of a gene of *Legionella pneumophila* Paris and/or Lens strain and/or Philadelphia strain, or of associated microorganisms. In effect, the presence of the expression product of a gene recognized by a specific antibody of said product expression can be detected by the presence of an antigen-antibody complex formed after contact of the strain of *Legionella pneumophila* Paris, Lens or Philadelphia strain, or of the associated microorganism with an antibody according to the invention. The bacterial strain utilized can were « prepared », that is, centrifuged, lysated, placed in an appropriate reagent for the constitution of the medium prone to immunological reaction. In particular, a detection process for expression in the gene is preferred, corresponding to a Western blot, capable of being carried out after electrophoresis on polyacrylamide gel of a lysate of the bacterial strain, in the presence or in the absence of reductory conditions (SDS-PAGE). After migration and separation of the proteins on the polyacrylamide gel, said proteins are transferred to an appropriate membrane (for example made of nylon) and the presence of the protein or of the polypeptide of interest is detected, by contact of said membrane with an antibody according to the invention.

Therefore, the present invention likewise comprises the kits or the necessary for implementing a process such as described (detection of the expression of a gene of *Legionella pneumophila* Paris and/or Lens strain and/or Philadelphia, or of an associated microorganism, or for detection and/or identification of bacteria belonging to

the *Legionella pneumophila* species Paris and/or Lens and/or Philadelphia strain, or an associated microorganism), comprising the following elements:

a) a polyclonal or monoclonal antibody according to the invention;

b) possibly, the reagents for the constitution of the medium prone to immunological reaction;

c) possibly, the reagents allowing use of the antigen-antibody complexes produced by the immunological reaction.

The polypeptides and the antibodies according to the invention can advantageously be immobilized on a support, especially a protein chip. This type of protein chip is an object of the invention, and can likewise contain at least one polypeptide of a microorganism other than *Legionella pneumophila* Paris and/or Lens strain and/or Philadelphia strain, or an antibody directed against a compound of a microorganism other than *Legionella pneumophila* Paris and/or Lens and/or Philadelphia strain.

The high-density protein or filter chips containing proteins according to the invention can be constructed in the same way as the DNA chips according to the invention. In practice, synthesis of the polypeptides fixed directly on the protein chip, or *ex situ* synthesis followed by a stage for fixing the synthesized polypeptide on said chip can be carried out. This latter method is preferable, when proteins of significant size are to be fixed on the support, which are advantageously prepared by genetic engineering. All the same, if it is preferred to fix only peptides on the support of said chip, it can be more interesting to proceed with synthesis of said peptides directly *in situ*.

The protein chips according to the invention can be advantageously utilized in kits or necessary for the detection and/or identification of bacteria associated with the *Legionella pneumophila* species Paris and/or Lens strain and/or Philadelphia strain, or with a microorganism, more generally in kits or necessary for the detection and/or identification of microorganisms. When the polypeptides according to the invention are fixed on DNA chips, the presence of antibodies is searched for in the samples tested, with the fixation of an antibody according to the invention on the support of the protein chip allowing identification of the protein whereof said antibody is specific.

Preferably, an antibody according to the invention is fixed on the support of the protein chip, and the presence of the corresponding antigen, specific to *Legionella pneumophila* Paris and/or Lens strain and/or Philadelphia strain or of an associated microorganism is detected.

A protein chip described hereinabove can be utilized for the detection of products of genes, for establishing an expression profile of said genes, as a complement to a DNA chip according to the invention.

The protein chips according to the invention are likewise extremely useful for proteomic experiments, which studies the interactions between the different proteins of a given microorganism. In a simplified manner, peptides representative of the different proteins of an organism are fixed on a support. Next, said support is put in contact with marked proteins, and after an optional rinsing stage, interactions are detected between said marked proteins and the peptides fixed on the protein chip.

Therefore, the protein chips comprising a polypeptidic sequence according to the invention or an antibody according to the invention are an object of the invention, as well as kits or necessary containing them.

The present invention likewise covers a process for detection and/or identification of bacteria belonging to the *Legionella pneumophila* species Paris and/or Lens and/or Philadelphia strain, or to an associated microorganism in a biological sample, which uses a nucleotidic sequence according to the invention.

It should be understood that the term biological sample relates in the present invention to the samples taken from a living organism (in particular blood, tissue, organs or the like taken a mammal) or a sample containing biological material, that is, DNA. Such a biological sample especially includes all fluids (liquid or air-borne), or any object, such as conduits for fluid, filters for fluids, or any object capable of being implied in the fluid supply in buildings, nutritional compositions containing bacteria or other.

The process for detection and/or identification utilizing the nucleotidic sequences according to the invention can be diverse in nature.

A process comprising the following stages is preferred:

a) possibly, isolation of DNA from the biological sample to be analyzed, or obtaining DNAc from the RNA of the biological sample;

b) specific amplification of the DNA of bacteria belonging to the *Legionella pneumophila* species Paris strain, Lens, Philadelphia or to a microorganism associated by means of at least one primer according to the invention;

c) revealing amplification products.

This process is based on specific amplification of DNA, in particular by a chain amplification reaction.

Likewise, a process comprising the following stages is preferred:

a) contact by a nucleotidic probe according to the invention with a biological sample, the nucleic acid contained in the biological sample having, if required, previously been made accessible to hybridization, in conditions permitting hybridization of the probe to the nucleic acid of bacteria belonging to the *Legionella pneumophila* species Paris strain, or to an associated microorganism;

b) revealing the hybrid possibly formed between the nucleotidic probe and the DNA of the biological sample.

Such a process must not be limited to detection of the presence of the DNA contained in the biological sample tested, and it can likewise be used for detecting the RNA contained in said sample. This process in particular includes the Southern and Northern blot.

Another preferred process according to the invention comprises the following stages:

a) contact by a nucleotidic probe immobilized on a support according to the invention with a biological sample, the nucleic acid of the sample, having, if required, been previously made accessible to hybridization, in conditions allowing hybridization of the probe to the nucleic acid of bacteria belonging to the *Legionella pneumophila* species Paris strain or to an associated microorganism;

b) contact by the hybrid formed between the nucleotidic probe immobilized on a support and the nucleic acid contained in the biological sample, if required after elimination of the DNA from the biological sample not having hybridized with the probe, with a nucleotidic probe marked according to the invention;

c) revealing the novel hybrid formed at stage b).

This process is advantageously utilized with a DNA chip according to the invention, the desired nucleic acid hybridizing with a probe present on the surface of said chip, and being detected by utilization of a marked probe. This process is advantageously implemented by combining a previous amplification stage of DNA or of complementary DNA obtained possibly by inverse transcription, by means of primers according to the invention.

Therefore, the present invention likewise includes the kits or necessary for the detection and/or identification of bacteria belonging to the *Legionella pneumophila* species Paris and/or Lens and/or Philadelphia strain, or to an associated microorganism, characterized in that it comprises the following elements:

a) a nucleotidic probe according to the invention;

b) possibly, the reagents necessary for using a hybridization reaction;

c) possibly, at least one primer according to the invention as well as the reagents necessary for amplification reaction of the DNA.

Similarly, the present invention likewise includes the kits or necessary for the detection and/or identification of bacteria belonging to the *Legionella pneumophila* species Paris strain or to an encircled microorganism, characterized in that it comprises the following elements:

a) a nucleotidic probe, known as capture probe, according to the invention;

b) an oligonucleotidic probe, known as revelation probe, according to the invention;

c) possibly, at least one primer according to the invention as well as the reagents necessary for amplification reaction of the DNA.

Finally, the kits or necessary for the detection and/or identification of bacteria belonging to the *Legionella pneumophila* species Paris and/or Lens and/or Philadelphia strain, or to an associated microorganism, characterized in that it comprises the following elements:

a) at least one primer according to the invention;

b) possibly, the reagents necessary for performing an amplification reaction of DNA;

c) possibly, a compound enabling the sequence of the amplified fragment, more particularly an oligonucleotidic probe according to the invention, to be verified,

are likewise objects of the present invention.

Preferably, said primers and/or probes and/or polypeptides and/or antibodies according to the present invention utilized in the processes and/or kits or necessary according to the present invention are selected from amongst the primers and/or probes and/or polypeptides and/or antibodies specific to the *Legionella pneumophila* species Paris and/or Lens and/or Philadelphia strain. In a preferred manner these elements are selected from amongst the nucleotidic sequences coding for a secreted protein, among the polypeptides secreted, or among the antibodies directed against exported polypeptides, such as those implied in the wall or the cellular envelope of *Legionella pneumophila* Paris and/or Lens and/or Philadelphia strain.

The object of the present invention is likewise the strains of *Legionella pneumophila* Paris or Lens strain, and/or associated microorganisms containing one or

more mutation(s), at most less than 10 % mutation (or again less as cited for the modifications of polypeptides) in a nucleotidic sequence according to the invention, in particular an ORF sequence, or their regulatory elements (in particular promoters).

According to the present invention, the strains of *Legionella pneumophila* Paris or Lens strain having one or more mutation(s) in the nucleotidic sequences coding for polypeptides implied in the machine cellular, in particular secretion, central intermediary metabolism, energetic metabolism, the process of synthesizing of the amino acids, transcription and translation, synthesis of the polypeptides, are preferred.

Said mutations can lead to inactivation of the gene, or in particular when they are situated in the regulatory elements of said gene, at overexpression of the latter.

The invention relates further to utilizing a compound selection method capable of inhibiting the expression of genes implied in the biosynthesis of polysaccharides having a cellular envelope of bacteria of the *Legionella pneumophila* species Paris strain, characterized in that it comprises the following stages:

a) contact by said compound with a bacteria of said Paris strain, said bacteria being in conditions and in medium appropriate to its culture;

b) determination of the capacity of said compound to inhibit the expression of the genes coding for the proteins of SEQ ID Nos. 1126, 3218, 288, 632, 917, 1503, 1555, 1877, 1928, 1963, 2204, 2212, 2243, 2324, 2378, 2410, 2411;

c) by means of a process according to the invention in which said antibody is directed specifically against a polypeptide implied in the biosynthesis of the polysaccharides, or by means of a process according to the invention in which the probes or primers are specific to a nucleic sequence coding for a polypeptide implied in the biosynthesis of the polysaccharides;

d) selection of organic or inorganic compound capable of modulating, regulating, inducing or inhibiting the expression of genes, and/or of modifying the cellular replication of eukaryotic or prokaryotic cells or capable of inducing, inhibiting or aggravating the pathologies associated with infection by *Legionella pneumophila* Paris strain or one of its associated microorganisms.

The invention likewise comprises a method for selection of compounds capable of binding to a polypeptide or one of its fragments according to the invention, capable of binding to a nucleotidic sequence according to the invention, or capable of recognizing an inventive antibody, and/or capable of modulating, regulating, inducing or inhibiting the expression of genes, and/or modifying the growth or cellular

replication of eukaryotic or prokaryotic cells, or capable of inducing, inhibiting or aggravating in a animal or human organism the pathologies bound to infection by *Legionella pneumophila* Paris strain or Lens or Philadelphia strains, or one of its associated microorganisms, characterized in that it comprises the following stages:

a) contact by said compound with said polypeptide, said nucleotidic sequence, with a cell transformed according to the invention and/or administration of said compound to an animal transformed according to the invention;

b) determination of the capacity of said compound to be bound with said polypeptide or said nucleotidic sequence, or to modulate, regulate, induce or inhibit the expression of genes, or to modulate the growth or cellular replication, or induce, inhibit or aggravate in said transformed animal the pathologies bound to infection by *Legionella pneumophila* Paris strain or Lens or Philadelphia strains strain, or one of its associated microorganisms.

The cells and/or the animals transformed according to the invention, could advantageously serve as a model and be utilized in processes for studying, identifying and/or selecting compounds capable of being responsible for pathologies induced or aggravated by *Legionella pneumophila* Paris strain or Lens or Philadelphia strains strain, or capable of preventing and/or treating these pathologies such as for example genital, ocular or systemic diseases, especially of the lymphatic system. In particular, the transformed host cells, especially the bacteria of the family of *Legionellae* whereof the transformation by a vector according to the invention can for example grow or inhibit its infectious capacity, or modulate the pathologies habitually induced or aggravated by the infection, could be utilized for infecting animals in which the appearance of pathologies will be followed. These animals not transformed, infected for example with transformed *Legionellae* bacteria, will be able to serve as a study model. In the same manner, the animals transformed according to the invention will be able to be utilized in selection processes for compounds capable of preventing and/or treating diseases due to *Legionella*. Said processes utilizing said transformed cells and/or transformed animals are part of the invention.

The compounds capable of being selected can be organic compounds such as polypeptides or hydrates of carbon or any other already known organic or inorganic compounds, or novel organic compounds elaborated by molecular modeling techniques and obtained by chemical or biochemical synthesis, these techniques being known to the specialist.

Said selected compounds will be able to be utilized for modulating the growth and/or cellular replication of *Legionella pneumophila* Paris and/or Lens and/or Philadelphia strain, or any other associated microorganism and thus to control infection by these microorganisms. Said compounds according to the invention will likewise be utilized for modulating the growth and/or cellular replication of all eukaryotic or prokaryotic cells, especially tumoral cells and infectious microorganisms, for which said compounds will prove to be active, the methods determining said modulations being well known to the specialist.

Compound capable of modulating the growth of a microorganism is understood to mean any compound allowing to intervene, modify, limit and/or reduce the development, growth, rate of proliferation and/or the viability of said microorganism.

This modulation can be realized for example by an agent capable of binding to a protein and thus inhibit or potentialize its biological activity, or capable of binding to a membranous protein of the external surface of a microorganism and blocking penetration of said microorganism in the host cell or benefiting the action of the immune system of the infected organism directed against said microorganism. This modulation can likewise be realized by an agent capable of binding to a nucleotidic sequence of DNA or RNA of a microorganism and for example blocking the expression of a polypeptide whereof the biological or structural activity is necessary to the growth or to the reproduction of said microorganism.

For these screening methods, likewise associated microorganism in the present selection method is understood to mean any microorganism whereof the gene expression can be modulated, regulated, induced or inhibited, or whereof the growth or cellular replication can be likewise modulated by a compound of the invention. Likewise, associated microorganism in the present invention is understood to mean any microorganism comprising nucleotidic sequences or polypeptides according to the invention. These microorganisms can in certain cases comprise polypeptides, or nucleotidic sequences identical or homologous to those of the invention will likewise be able to be detected and/or identified by the processes or detection and/or identification kit according to the invention and likewise serve as target for the compounds of the invention.

The invention relates to the compounds capable of being selected by a selection method according to the invention.

The invention likewise relates to a pharmaceutical composition comprising a compound selected from amongst the following compounds:

a) a nucleotidic sequence according to the invention;

b) a polypeptide according to the invention;

c) a vector according to the invention;

d) an antibody according to the invention; and

e) a compound capable of being selected by a selection method according to the invention, possibly in association with a pharmaceutically acceptable vehicle.

Efficacious quantity is understood to mean an adequate quantity of said compound or antibodies, or of polypeptide of the invention, allowing to modulate the growth of *Legionella pneumophila* Paris and/or Lens and/or Philadelphia strain, or of an associated microorganism.

The invention also relates to a pharmaceutical composition according to the invention for the prevention or treatment of an infection by a bacteria belonging to the *Legionella pneumophila* species Paris strain or Lens or Philadelphia strains, or by an associated microorganism.

The further aim of the invention is an immunogenic and/or vaccinal composition, characterized in that it comprises one or more polypeptides according to the invention and/or one or more hybrid polypeptides according to the invention.

The invention also comprises utilization of a cell transformed according to the invention, for the preparation of a vaccinal composition.

The aim of the invention likewise is a vaccinal composition, characterized in that it contains a nucleotidic sequence according to the invention, a vector according to the invention and/or a cell transformed according to the invention.

The invention likewise relates to vaccinal compositions according to the invention, for the prevention or treatment of an infection by a bacteria belonging to the *Legionella pneumophila* species Paris strain or Lens or Philadelphia strains, or by an associated microorganism.

In a preferred manner the immunogenic and/or vaccinal compositions according to the invention for preventing and/or treating infection by *Legionella pneumophila* Paris strain or Lens or Philadelphia strains, or by an associated microorganism will be selected from among the immunogenic and/or vaccinal compositions comprising a polypeptide or one of its fragments corresponding to a protein, or one of its fragments, of the cellular envelope of *Legionella pneumophila* Paris strain or Lens or Philadelphia

strains. The vaccinal compositions comprising nucleotidic sequences will preferably likewise comprise nucleotidic sequences coding for a polypeptide or one of its fragments corresponding to a protein, or one of its fragments, of the cellular envelope of *Legionella pneumophila* Paris strain or Lens or Philadelphia strains.

Of these preferred immunogenic and/or vaccinal compositions, the most preferred are those comprising a polypeptide or one of its fragments, or a nucleotidic sequence or one of its fragments whereof the sequences are selected from among the nucleotidic sequences or amino acids identified in this functional group and listed previously.

The polypeptides of the invention or their fragments entering the immunogenic compositions according to the invention can be selected by techniques known to the specialist such as for example on the capacity of said polypeptides to stimulate the T cells, which translates for example by their proliferation or the secretion of interleukins, and which terminates with the production of antibodies directed against said polypeptides.

In mice, in which a ponderal dose of the vaccinal composition comparable to the dose utilized in humans is administered, the reaction antibody is tested by taking serum followed by studying the formation of a complex between the antibodies present in the serum and the antigen of the vaccinal composition, according to customary techniques.

According to the present invention, said vaccinal compositions will preferably be in association with a pharmaceutically acceptable vehicle and, if required, with one or more adjuvants of appropriate immunity.

These days, diverse types of vaccines are available for protecting humans against infectious diseases: attenuated living microorganisms (*M. bovis* - BCG for tuberculosis), inactive microorganisms (flu virus), acellular extracts (*Bordetella pertussis* for pertussis), recombinant proteins (surface antigen of the hepatitis B virus), polyosides (pneumococci). Vaccines prepared from synthesis peptides or genetically modified microorganisms expressing heterologous antigens are in experimentation. Still more recently, recombined plasmidic DNAs carrying coding for protector antigens were proposed as an alternative vaccinal strategy. This type of vaccination is realized with a particular plasmid deriving from a plasmid of *E. coli* which does not replicate *in vivo* and which codes solely for the vaccinating protein. Animals were immunized by simply injecting naked plasmidic DNA into the muscle. This technique results in expression of the vaccinal protein *in situ* and to an immune response of cellular type (CTL) and of

humoral type (antibodies). This double induction of the immune response is one of the principal advantages of the vaccination technique with naked DNA.

The vaccinal compositions comprising nucleotidic sequences or vectors in which said sequences are inserted are described in particular in international application N° WO 90/11092 and likewise in international application N° WO 95/11307.

The nucleotidic sequence making up the vaccinal composition according to the invention can be injected into the host after having been coupled to compounds which benefit penetration of this polynucleotide inside the cell or its transport as far as the cellular nucleus. The resulting conjugates can be encapsulated in polymer microparticles, as described in international application N° WO 94/27238 (Medisorb Technologies International).

According to another embodiment of the vaccinal composition according to the invention, the nucleotidic sequence, preferably a DNA, is complexed with DEAE-dextran, with nuclear proteins, with lipids or encapsulated in liposomes or introduced in the form of a gel facilitating its transfection in the cells. The polynucleotide or the vector according to the invention can also be in suspension in a buffer solution or be associated with liposomes.

Advantageously, such a vaccine will be prepared according to the technique described by Tacson *et al.* or Huygen *et al.* in 1996 or again according to the technique described by Davis *et al.* in the international application N° WO 95/11307.

Such a vaccine can likewise be prepared in the form of a composition containing a vector according to the invention, placed under the control of regulation elements allowing its expression in humans or animals. For example, the polypeptidic antigen of interest, the pcDNA3 plasmid or the pcDNA1/neo plasmid could be utilized as an *in vivo* expression vector, both marketed by Invitrogen (R & D Systems, Abingdon, UK). Such a vaccine will comprise advantageously, apart from the recombinant vector, a saline solution, for example a sodium chloride solution.

Pharmaceutically acceptable vehicle is understood to mean a compound or a combination of compounds entering a pharmaceutical or vaccinal composition not causing secondary reactions and which enables for example ease of administration of active compound, an increase in its life expectancy and/or its efficacy in the organism, augmentation of its solubility in solution or again an improvement in its preservation. These pharmaceutically acceptable vehicles are well known and will be adapted by the

specialist as a function of the nature and mode of administration of the selected active compound.

As for vaccinal formulations, these can comprise adjuvants of appropriate immunity which are known to the specialist, such as for example aluminum hydroxide, a representative of the family of muramyl peptides as one of the peptidic derivatives of N-acetyl-muramyl, a bacterial lysate, or even the incomplete Freund adjuvant.

Preferably, these compounds will be administered systemically, in particular intravenously, intramuscularly, intradermally or subcutaneously, or orally. In a more preferred way, the vaccinal composition comprising polypeptides according to the invention will be administered in several doses, spread out over time, intradermically or subcutaneously.

Their administration methods, posologies and optimal galenic forms can be determined according to the criteria generally taken into account in setting up treatment adapted to a patient such as for example age or body weight of the patient, the seriousness of the general status, tolerance to treatment and the secondary effects.

The invention comprises utilizing an inventive composition, for the treatment or prevention of diseases brought on or aggravated by the presence of *Legionella pneumophila* Paris strain or Lens or Philadelphia strains.

The invention comprises the utilization of a composition according to the invention for the treatment or prevention of systemic diseases, induced or aggravated by the presence of *Legionella pneumophila* Paris strain or Lens or Philadelphia strains.

Additionally, an object of the present invention likewise is a genomic DNA bank of a bacteria of the species *Legionella pneumophila* Paris strain, characterized in that this is the bank deposited with the CNCM on November 19, 2003, under the order number I-3138.

Additionally, an object of the present invention likewise is a genomic DNA bank of a bacteria of the species *Legionella pneumophila* Lens strain, characterized in that this is the bank deposited with the CNCM on September 23, 2004, under the order number I-3306.

Additionally, an object of the present invention likewise is a vector or a host cell as claimed in Claim 38 or 42, characterized in that this is the vector or the cell deposited with the CNCM on November 19, 2003, under the order number I-3137.

One of the advantages of using the BAC system relative to a cosmides system is that the plasmid utilized is present only in a maximum two copies per transformed cell,

which reduces the potential for recombination between DNA fragments, and more importantly, which eliminates the risk of lethal overexpression of bacterial cloned genes. Nevertheless, the presence of BAC as a single copy signifies that the plasmidic DNA must be extracted from a large volume of culture in order to obtain enough DNA for the sequence. In addition, the stability and fidelity of maintenance of the clones in a BAC bank enable identification of genomic differences among different strains of *Legionella*, and identification of these genetic differences which can be responsible for the phenotypical variations observed between the different strains.

The genomic DNA banks described in the present invention effectively cover the genome of *Legionella pneumophila* Paris and Lens strains. All the same, although it is possible that certain regions have not been able to be cloned in said bank, by virtue of lethality problems in *Escherichia coli,* these regions can easily be amplified and identified by the specialist, by utilizing oligonucleotides specific to the sequences of the ends of the different clones which form the contigs.

Additionally, an object of the present invention likewise is a method for isolating a polynucleotide of interest present in a bacteria of the *Legionella* genre and absent from a bacteria of another genre, or present in a pathogenic bacteria of the *Legionella* genre and absent from a non-pathogenic bacteria of the *Legionella* genre, or again present in a bacteria of the *Legionella pneumophila* species and absent from a bacteria of any other species of the *Legionella* genre, or again present in a bacteria of the *Legionella pneumophila* species Paris and/or Lens and/or Philadelphia strain and absent from a bacteria of the *Legionella pneumophila* species of any other strain, characterized in that it utilizes at least the BAC bank deposited on November 19, 2003 (I-3138) with the C.N.C.M and the BAC bank deposited on September 23, 2004 (I-3306) with the CNCM according to the invention.

Said method is preferably characterized in that it comprises the following stages:

a) isolating at least one polynucleotide contained in a clone of said DNA bank deposited with the CNCM on November 19 2003, under the order number I-3138 or contained in a clone of said BAC bank deposited on September 23 2004 under the number I-3306;

b) isolating:

- at least one genomic polynucleotide or DNAc of a second bacteria of another genre or of the *Legionella* genre, said second bacteria of the *Legionella* genre belonging to a different strain of the Paris strain or, alternatively,

- at least one polynucleotide contained in a clone of a DNA bank based on a BAC prepared from the genome of a second bacteria of another genre or of the *Legionella* genre, said second bacteria of the *Legionella* genre belonging to a different strain of the Paris and/or Lens and/or Philadelphia strain;

c) hybridizing the polynucleotide of stage a) to the polynucleotide of stage b);

d) selecting the polynucleotides of stage a) which do not have the hybridization complex form with the polynucleotides of stage b); and

e) characterizing the selected polynucleotide.

The polynucleotide of stage a) can be prepared by the digestion of at least one recombinant BAC clone with an appropriate restriction enzyme, and optionally, amplification of the resulting polynucleotide insert.

Therefore, the method of the invention enables the specialist to effect comparative genomic studies between the different strains or species of the Legionella genre, for example between the pathogenic strains and their non-pathogenic equivalent.

In particular, it is possible to study and determine the regions of polymorphism between said strains.

LEGENDS OF THE FIGURES

Figure 1: Circular genomic map of the line L. pneumophila Paris and specific genes of the L. pneumophila Lens line. From the exterior: circle 1: genes of Paris line on the chains + and – respectively. Red line, inversion in line Lens. Color code: green, genes of Paris line, black, rRNA operons, red, known virulence genes; the numbers indicate their position: 1 lvh-lvr secretion system type IV (lvrABC, lvhB2B3B4B5, lvrD, lvhB6B8B9B10B11D4, lvrE); 2 dot/icm secretion system type IV (icmTSRQOMLKEGCDJBF); 3 mip, 4 lspA, 5 lspDE, 6 htrA, 7 lspFGHIJK, 8 enhABC, 9 dot/icm secretion system type IV (icmVWX and dotABCD), 10 momp; circle 2: specific genes of the Lens line relative to the Paris line; 3: bias G/C (G+C/G-C) of the Paris line; circle 4: G+C content of the Paris line with <32,5% G+C in light yellow, between 32.5% and 44.1% in yellow and with >44.1% G+C in dark yellow. The scale (Mb) is indicated on the outside, the origin of the replication in position 0.

Figure 2: Phylogenetic tree of a multiple sequential comparison of kinase domains from Legionella pneumophila of Paris line to other prokaryotic and eukaryotic kinases by utilizing the MEGA program. The calculation was made by utilizing the

Poisson correction as a distance process and as a tree construction process the Neighbor-joining distance method. 0.2 indicates the 2 amino acid substitutions for 10 sites. The Nrprot access numbers, the names of genes and the names of organisms are indicated on the pattern. The numbers indicate the priming values.

Figure 3: Schema representative the genome core and the sole gene complement of L. pneumophila Paris, Lens and Philadelphia lines. The orthologous genes were defined by the most adequate reciprocated FASTA comparisons. The threshold was defined at a maximum of 80 % of sequence identities and at a length ratio of 0.75 to 1.33. The coding sequences of Philadelphia were determined by the Genmark predictions utilizing the « CAAT-box » program and the sequence obtained on the site http://www.genome3.cpmc.columbia.edu/-legion/project/ (latest version).

Figure 4: A. Comparison of the protein-coding RTX genes of L. pneumophila AA100, Paris and Lens lines. The sequence of the rtxA locus of line AA100 was obtained from the NCBI database (AAD41583). The dotted lines indicate that the correct number of repetitions is uncertain. B. Consensus sequences of the highly preserved repeated patterns of Paris and Lens lines. The amino-acid sequences in black indicate 11 amino-acids of the preserved N-terminal sequence of Paris and Lens lines, the amino-acids sequences in color represent the repeated patterns of each line (same color as for A). The underlined amino-acids indicate the positions which can change among the repetitions.

Figure 5: Pattern illustration of the different stages of intracellular growth of L. pneumophila in the macrophages. The different phases are numbered 1) Adhesion and invasion of L. pneumophila in the host cell 2) The phagosome does not fuse with the lysosomes but recruits organelles and converts to a compartment of rugged endoplasmic reticulum type. 3) Intracellular replication, non-flagellated L. pneumophila inside a phagosome. 4) Release of L. pneumophila. Flagellated. In red: Different important stages in the infectious cycle of L. pneumophila. In blue: Hypothesis indicating the stages at which the identified proteins could interfere in this cycle.

Figures 6 and 7: Southern Blot showing the specificity of the repeated sequence SEQ ID N° 7074 in L. pneumophila; the legend of Figure 6 given by Table XXV and that of Figure 7 by Table XXVI.

EXAMPLES

Example 1 : Materials and methods

1. Construction of banks

Shotgun bank of small fragments (size 1.5 to 2.5 kb)

The chromosomal DNA of the strains studied was prepared by a classic method including proteinase K treatment and phenol extraction (9). Approximately 36 ug of DNA were broken by nebulization (1 minute under pressure of 1 bar) (4). The ends of the DNA fragments were rendered free by having the DNA-polymerase of the bacteriophage T4 act for 15 minutes at 37°C in the presence of 4 tri-phosphate nucleotides. The enzyme was inactivated by incubation of 15 mn at 75°C. Adaptors (invitrogen Cat. N° 408-18) were ligatured to these ends. After ligature, the fragments of chromosomal DNA of a size between 1500 and 2500 base pairs were purified after electrophoresis on agarose gel. The vector utilized for construction of the bank, pcDNA2.1 (Invitrogen), was digested by the BstX1 enzyme and purified by geneclean (BIO-101) after electrophoresis on agarose gel. The chromosomal DNA and the purified vector were ligatured by action of the ligase of the bacteriophage T4. The ligation mixture was introduced by transformation to the strain of *Escherichia coli* XL2-blue (Stratagene). Environ 4000 colonies are obtained per ul of the ligation mixture.

Bank of average fragments (size 5 to 10 kb)

The bank was constructed by the technique of 'partial fill in' in the vector pSYX34 (12). The chromosomal DNA of the strain *L. pneumophila* Paris was prepared by partial digestion by *Sau*IIIA (Roche). After precipitation of the DNA in sodium acetate and the stage of partial fill-in with the A and G nucleotides by utilizing the Klenow enzyme, the fragments of chromosomal DNA having a size of between 5000 and 10000 base pairs were purified after electrophoresis on agarose gel and geneclean.

The vector is prepared in the same way by partial digestion with the *Sal*I enzyme, precipitation in sodium acetate then reaction of partial fill-in with the C and T nucleotides and purification on agarose gel and geneclean. The fragments of chromosomal DNA and the purified vector were ligatured by action of the ligase of the bacteriophage T4. The ligation mixture was introduced by transformation to the strain of *Escherichia coli* XL10-Gold (Stratagene). Around 4000 colonies are obtained per ul of the ligation mixture. The two ends of around 4000 fragments of this bank were sequenced.

Bank of large fragments (size 25 to 90 kb)

The bank of large fragments was constructed as described previously (4) by utilizing the pIndigo BAC vector (Epicentre). Briefly, in order to avoid mechanical

breaking of the DNA molecules the cells were included in agarose blocks in which DNA extraction is performed directly. For the preparation of large-sized fragments we performed partial digestion by *Hind*III (Roche) and separation by electrophoresis in pulsed fields. Fragments of sizes between 40 and 80 and between 80 and 130 kb were excised from the gel, purified by agarase treatment and ligature with the vector. The ligation mixture was introduced by electroporation to the strain of *Escherichia coli* DH10B (Gibco BRL). 1300 colonies were stored. The plasmidic DNAs of these 1300 colonies were extracted and the two ends of the cloned fragments were sequenced.

2. Preparation of plasmids and sequencing.

The plasmids were prepared by a semi-automatic preparation method developed at the GMP laboratory and based on the alkaline lysis method (2). The chromosomal inserts were sequenced from their two ends by utilizing the T7 and universal primer in following the recommendations of the supplier (Applied-Biosystems). The sequences were determined by utilizing automatic sequencers of type 3700 (Applied-Biosystem).

3. Assembling of sequences.

The sequences were assembled by utilizing the software suite developed at the University of Washington, Phred, Phrap and Consed (5, 8). The sequence completion was done by utilizing the software suite CAAT-box (7). The finishing stage corresponds to resequencing of the regions where the sequence is only slightly secure and sequencing of the regions located between the contigs. It was carried out either by sequencing PCR products or by operating on the bank clones. The sequences of oligonucleotides were defined by utilizing the consed and Primo software (8, 10).

4. Annotation of sequences.

The identification of the phases coding (CDS) was done by utilizing the software suite CAAT-box (7). This program combines the results of different methods:

- (i) identification of open reading phases and their tri as a function of their size;
- (ii) analysis of the probability of being coding by utilizing the Genemark software (11);
- (iii) identification of translation start (initiation codon and fixing sequence of the ribosome); and
- (iv) the % of identity of the proteic sequence deduced with the proteic sequences contained in sequence banks by utilizing BLASTP software.

The functions of the proteins coded by the identified coding phases were predicted by analysis of the search results of similarities in the non-redundant NCBI bank (http://www.ncbi.nlm.nih.gov/BLAST/) by utilizing BLASTP software (1).

5. Comparison of the genomes – identification of the CDS specific to the strain of *L. pneumophila* Paris strain.

All the proteic sequences deduced from the predicted coding phases of each genome was compared to all the proteic sequences possibly coded by the other genome by using BLASTP software. A threshold of 75 % of identity on the totality of the length of the protein was retained to identify the specific proteins of an isolate. This very high value was retained since it best allows discrimination of the orthologous genes from the paralogous genes (6). For the proteic sequences for which the sequence preservation is high (> at 70 %) the preservation of the nucleotidic sequences of the genes will also be high and could give a signal in low-stringency hybridization conditions. It will be necessary to consider this eventuality in the analysis of the test result.

Example 2: Deposit of biological material

The following organisms were deposited on November 19, 2003 with the Collection Nationale de Cultures de Microorganisms (CNCM) [National Collection of Microorganism Cultures], 25 rue du Docteur Roux, 75724 Paris Cedex 15, France, according to the dispositions of the Budapest Treaty:

- Clone of a shotgun bank, clone in the pCDNA vector, of the genome of *Legionella pneumophila* Paris strain (Pasteur Institute Collection CIP 107-629-T), registered under the file number I-3137. The insert of this clone is at a size of 14.2 kb and contains a gene coding for an autotransporter called led0019A07;
- BAC DNA bank (1248 clones) of the genome of *Legionella pneumophila* Paris strain (Pasteur Institute Collection CIP 107-629-T), registered under file number I-3138. Said bank BAC (I-3138) was made in the *E.coli* DH10B strain (Grant *et al.*, PNAS, 87:4645, 1990). The inserts of this bank were cloned in the pBelo BAC-Kan vector (Mozo *et al.*, Mol. Gen. Genet., 1998, 258:562-70) and have an average size of between 1.5 and 2.5 kb. The total of these inserts corresponds to complete coverage of the genome.

Example 3: Annotations of sequences

1. Genes specific to *L. pneumophila* Paris strain relative to the *L. pneumophila* Philadelphia strain

No significant identity between the nucleotidic sequence of the gene of *L. pneumophila* Paris strain and the genome of *L. pneumophila* Philadelphia strain.

Table VII: Example of annotation of sequences in the case of proteic and nucleic sequence of *L. pneumophila* Paris strain not having % of significant identity with respectively proteic and nucleic sequences of *L. pneumophila* Philadelphia strain

IPF No. of the gene of *L. pneumophila* Paris strain	IPF No. of the gene of *L. pneumophila* Philadelphia strain (best score)	% of identity of proteic sequences	% of identity of nucleotidic sequences
2043.1	-	-	-
2094.2	-	-	-
2039.1	-	-	-
2051.2	3061.1	33 %	not significant
3425.1	5305.1	32 %	not significant

2. Genes common to the two strains *L. pneumophila* Paris strain and Philadelphia strain for which the % of identity of deduced nucleic and proteic sequences is less than 75 %

Table VIII: Example of annotation of sequences in the case of proteic and nucleic sequences of genes common to the two strains of *L. pneumophila* Paris strain and Philadelphia strain for which the % of identity of the deduced nucleic and proteic sequences is less than 75 %

IPF No. of the gene of *L. pneumophila* Paris strain	IPF No. of the gene of *L. pneumophila* Philadelphia strain (best score)	% of identity of proteic sequences	% of identity of nucleotidic sequences
2244.2	3793.1	63 %	59 %
258.2	1342.1	60 %	59 %

3. Genes common to *L. pneumophila* Paris strain and Philadelphia strain for which the % of identity of the deduced nucleic and proteic sequences is greater than 75 %

Table IX: Example of annotation of sequences in the case of proteic and nucleic sequences of genes common to the two strains of *L. pneumophila* Paris strain and Philadelphia strain for which the % of identity of the deduced nucleic and proteic sequences is greater than 75 %

IPF No. of *L. pneumophila* Paris strain gene	IPF No. of *L. pneumophila* Philadelphia strain gene (best score)	% of identity of proteic sequences	% of identity of nucleotidic sequences
4629.2	133.1	100 %	100 %
6079.1	4147.1	90 %	88 %

Example 4: Example of alignment of sequences

Presented hereinbelow are the alignments of sequences preserved in the Paris and Philadelphia strains. For each of the six examples which follow, we present an alignment of the nucleotidic sequences as well as alignment of the sequences of amino acids. The alignment of the sequences of amino acids is obtained by aligning the translated sequence of the ORF present in the Paris strain with the sequence originating from translation in the six phases of the contigs of the Philadelphia sequence. The sequence homology of these ORFs present in the two strains is very strong, as much in amino acids as in nucleotides.

```
TBLASTN 2.2.6 [Apr-09-2003]

Query= 1764.3 CONTIG=Contig42 POSCDS1=13736 POSCDS2=14869 SENS=p, Seq Id : 555
         (216 letters)

Database: contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.done

                                                                    Score    E
Sequences producing significant alignments:                         (bits) Value

LpPhiladelphia_Contig49                                             433     e-123

>LpPhiladelphia_Contig49
          Length = 376826

 Score =  433 bits (1114), Expect = e-123
 Identities = 215/215 (100%), Positives = 215/215 (100%)
 Frame = -2

Query: 1      HLALFDETYIKTILILLSICVLKKFILRYDMILNDIVLYDNFFMTFDYKDNFMSKGPYQS     60
               HLALFDETYIKTILILLSICVLKKFILRYDMILNDIVLYDNFFMTFDYKDNFMSKGPYQS
Sbjct: 297565 HLALFDETYIKTILILLSICVLKKFILRYDMILNDIVLYDNFFMTFDYKDNFMSKGPYQS 297386

Query: 61     FANRLISALKDRGYTASRSPNGICIKTLAEFTGASEQICRRYIRGDALPDYEKVKQLAFH    120
               FANRLISALKDRGYTASRSPNGICIKTLAEFTGASEQICRRYIRGDALPDYEKVKQLAFH
Sbjct: 297385 FANRLISALKDRGYTASRSPNGICIKTLAEFTGASEQICRRYIRGDALPDYEKVKQLAFH 297206

Query: 121    LQVNPGWLLFGEDENATTKKNEVDEKLLHYILKQSHHLYPISQGSNDDYADFVLGLIKEV    180
               LQVNPGWLLFGEDENATTKKNEVDEKLLHYILKQSHHLYPISQGSNDDYADFVLGLIKEV
Sbjct: 297205 LQVNPGWLLFGEDENATTKKNEVDEKLLHYILKQSHHLYPISQGSNDDYADFVLGLIKEV 297026

Query: 181    KAIDTSENNLLKIIDLAIGSISSYEEKRKKHSHAV    215
               KAIDTSENNLLKIIDLAIGSISSYEEKRKKHSHAV
Sbjct: 297025 KAIDTSENNLLKIIDLAIGSISSYEEKRKKHSHAV 296921


  Database: contigsLpPhiladelphia
    Posted date:  Nov 20, 2003 10:38 AM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
   0.321    0.139    0.402

Gapped
Lambda     K      H
   0.267   0.0410    0.140

Matrix: BLOSUM62

Query= 1764.3 CONTIG=Contig42 POSCDS1=13736 POSCDS2=14869 SENS=p
         (1134 letters)

Database: /home/Gmp/rusniok/projets/legionella/pourBrevet-
191103/contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.....................................................done

                                                                    Score    E
Sequences producing significant alignments:                         (bits) Value
```

```
LpPhiladelphia_Contig35                                                 2248  0.0

>LpPhiladelphia_Contig35
          Length = 48622

 Score = 2248 bits (1134), Expect = 0.0
 Identities = 1134/1134 (100%)
 Strand = Plus / Plus


Query: 1    atgatcagaaaaataatttatgttacaggtactcgtgccgattatggactgatgagagaa   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7091 atgatcagaaaaataatttatgttacaggtactcgtgccgattatggactgatgagagaa 7150


Query: 61   gtactaaaaagattacaccagtcagaagacattgacttatcgatttgtgtcactggtatg  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7151 gtactaaaaagattacaccagtcagaagacattgacttatcgatttgtgtcactggtatg 7210


Query: 121  catcttgatgctttgtatggaaatacagttaacgaaattaaagcagatcagttctcaata  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7211 catcttgatgctttgtatggaaatacagttaacgaaattaaagcagatcagttctcaata 7270


Query: 181  tgcggcattattcctgttgatcttgccaatgctcagcatagttctatggcaaaagctatc  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7271 tgcggcattattcctgttgatcttgccaatgctcagcatagttctatggcaaaagctatc 7330


Query: 241  ggccatgaacttttgggattcaccgaggtattcgaaagtgaaactcctgatgtcgtttta  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7331 ggccatgaacttttgggattcaccgaggtattcgaaagtgaaactcctgatgtcgtttta 7390


Query: 301  ttgctgggagatcgaggagaaatgcttgctgcggccatagcagcgatacatttaaatatc  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7391 ttgctgggagatcgaggagaaatgcttgctgcggccatagcagcgatacatttaaatatc 7450


Query: 361  ccggttgtacatctgcacggaggagagcgctctggaaccgttgatgaaatggtaaggcat  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7451 ccggttgtacatctgcacggaggagagcgctctggaaccgttgatgaaatggtaaggcat 7510


Query: 421  gcgatttccaaattatctcattatcattttgtcgcaacagaggcatccaaacaacgattg  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7511 gcgatttccaaattatctcattatcattttgtcgcaacagaggcatccaaacaacgattg 7570


Query: 481  attagaatgggtgagaaagaagaaaccatttttcaggttggtgctccaggcttggatgaa  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7571 attagaatgggtgagaaagaagaaaccatttttcaggttggtgctccaggcttggatgaa 7630


Query: 541  atcatgcagtataaaacgtctacacgtgatgtgtttaatcaacgttatggatttgatcct  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7631 atcatgcagtataaaacgtctacacgtgatgtgtttaatcaacgttatggatttgatcct 7690


Query: 601  gacaaaaaaatctgtttattaatctatcacccggttgttcaagaagttgactcgattaaa  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7691 gacaaaaaaatctgtttattaatctatcacccggttgttcaagaagttgactcgattaaa 7750
```

```
Query: 661  attcaatttcaaagcgtgattcaggcagcactcgctacaaatttacagattatttgcctt  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7751 attcaatttcaaagcgtgattcaggcagcactcgctacaaatttacagattatttgcctt 7810

Query: 721  gagcctaattccgatacgggtggtcatttaattcgagaagtgattcaggaatatattgat  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7811 gagcctaattccgatacgggtggtcatttaattcgagaagtgattcaggaatatattgat 7870

Query: 781  catcctgatgttagaattatcaagcacttacatcgtccggaatttattgattgtcttgca  840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7871 catcctgatgttagaattatcaagcacttacatcgtccggaatttattgattgtcttgca 7930

Query: 841  aattctgatgtgatgctgggaaattccagtagtggcatcatagaggcagcctcatttaac  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7931 aattctgatgtgatgctgggaaattccagtagtggcatcatagaggcagcctcatttaac 7990

Query: 901  ctgaacgtagttaatgttggaagcaggcaaaatttaagagaacgaagcgacaatgtcatt  960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 7991 ctgaacgtagttaatgttggaagcaggcaaaatttaagagaacgaagcgacaatgtcatt 8050

Query: 961  gatgttgatgttacttatgatgctattttgactggtctaagagaagcgctaaataaaccc 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 8051 gatgttgatgttacttatgatgctattttgactggtctaagagaagcgctaaataaaccc 8110

Query: 1021 aagataaaatactctaactgttatggggatggaaaaacgagtgaaaggtgttatcaattg 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 8111 aagataaaatactctaactgttatggggatggaaaaacgagtgaaaggtgttatcaattg 8170

Query: 1081 ttaaaaactatccctttgcactcacaaatattgaataaatgcaatgcatactaa       1134
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 8171 ttaaaaactatccctttgcactcacaaatattgaataaatgcaatgcatactaa       8224
```

```
TBLASTN 2.2.6 [Apr-09-2003]

Query= 1864.3 CONTIG=Contig42 POSCDS1=77740 POSCDS2=79155 SENS=p, Seq
Id : 622
         (489 letters)

Database: contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.done

                                                                  Score    E
Sequences producing significant alignments:                       (bits) Value

LpPhiladelphia_Contig49                                             1003   0.0

>LpPhiladelphia_Contig49
          Length = 376826

 Score = 1003 bits (2594), Expect = 0.0
 Identities = 488/488 (100%), Positives = 488/488 (100%)
```

```
 Frame = +2

Query: 1     KLSLPLIRLWQLSRSKHMFKPQGLYDYICQQWQEEILPSLCDYIKIPNKSPHFDAKWEEH    60
              KLSLPLIRLWQLSRSKHMFKPQGLYDYICQQWQEEILPSLCDYIKIPNKSPHFDAKWEEH
Sbjct: 21029 KLSLPLIRLWQLSRSKHMFKPQGLYDYICQQWQEEILPSLCDYIKIPNKSPHFDAKWEEH 21208

Query: 61    GYMEQAVNHIANWCKSHAPKGMTLEIVRLKNRTPLLFMEIPGQIDDTVLLYGHLDKQPEM   120
              GYMEQAVNHIANWCKSHAPKGMTLEIVRLKNRTPLLFMEIPGQIDDTVLLYGHLDKQPEM
Sbjct: 21209 GYMEQAVNHIANWCKSHAPKGMTLEIVRLKNRTPLLFMEIPGQIDDTVLLYGHLDKQPEM 21388

Query: 121   SGWSDDLHPWKPVLKNGLLYGRGGADDGYSAYASLTAIRALEQQGLPYPRCILIIEACEE   180
              SGWSDDLHPWKPVLKNGLLYGRGGADDGYSAYASLTAIRALEQQGLPYPRCILIIEACEE
Sbjct: 21389 SGWSDDLHPWKPVLKNGLLYGRGGADDGYSAYASLTAIRALEQQGLPYPRCILIIEACEE 21568

Query: 181   SGSYDLPFYIELLKERIGKPSLVICLDSGAGNYEQLWMTTSLRGNLVGKLTVELINEGVH   240
              SGSYDLPFYIELLKERIGKPSLVICLDSGAGNYEQLWMTTSLRGNLVGKLTVELINEGVH
Sbjct: 21569 SGSYDLPFYIELLKERIGKPSLVICLDSGAGNYEQLWMTTSLRGNLVGKLTVELINEGVH 21748

Query: 241   SGSASGIVADSFRVARQLISRIEDENTGEIKLPQLYCDIPDERIKQAKQCAEILGEQVYS   300
              SGSASGIVADSFRVARQLISRIEDENTGEIKLPQLYCDIPDERIKQAKQCAEILGEQVYS
Sbjct: 21749 SGSASGIVADSFRVARQLISRIEDENTGEIKLPQLYCDIPDERIKQAKQCAEILGEQVYS 21928

Query: 301   EFPWIDSAKPVIQDKQQLILNRTWRPALTVTGADGFPAIADAGNVMRPVTSLKLSMRLPP   360
              EFPWIDSAKPVIQDKQQLILNRTWRPALTVTGADGFPAIADAGNVMRPVTSLKLSMRLPP
Sbjct: 21929 EFPWIDSAKPVIQDKQQLILNRTWRPALTVTGADGFPAIADAGNVMRPVTSLKLSMRLPP 22108

Query: 361   LVDPEAASVAMEKALTQNPPYNAKVDFKIQNGGSKGWNAPLLSDWLAKAASEASMTYYDK   420
              LVDPEAASVAMEKALTQNPPYNAKVDFKIQNGGSKGWNAPLLSDWLAKAASEASMTYYDK
Sbjct: 22109 LVDPEAASVAMEKALTQNPPYNAKVDFKIQNGGSKGWNAPLLSDWLAKAASEASMTYYDK 22288

Query: 421   PAAYMGEGGTIPFMSMLGEQFPKAQFMITGVLGPHSNAHGPNEFLHLDMVKKLTSCVSYV   480
              PAAYMGEGGTIPFMSMLGEQFPKAQFMITGVLGPHSNAHGPNEFLHLDMVKKLTSCVSYV
Sbjct: 22289 PAAYMGEGGTIPFMSMLGEQFPKAQFMITGVLGPHSNAHGPNEFLHLDMVKKLTSCVSYV 22468

Query: 481   LYSFSQKK   488
              LYSFSQKK
Sbjct: 22469 LYSFSQKK 22492


  Database: contigsLpPhiladelphia
    Posted date:  Nov 20, 2003 10:38 AM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
   0.318    0.136    0.419

Gapped
Lambda     K      H
   0.267   0.0410    0.140

Matrix: BLOSUM62

Query= 1864.3 CONTIG=Contig42 POSCDS1=77740 POSCDS2=79155 SENS=p
         (1416 letters)

Database: /home/Gmp/rusniok/projets/legionella/pourBrevet-
191103/contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.....................................................done

                                                                  Score     E
Sequences producing significant alignments:                       (bits)  Value
```

```
LpPhiladelphia_Contig49                                          2807  0.0

>LpPhiladelphia_Contig49
          Length = 376826

 Score = 2807 bits (1416), Expect = 0.0
 Identities = 1416/1416 (100%)
 Strand = Plus / Plus


Query: 1     atgttcaaaccccaaggattgtatgattacatatgccaacagtggcaagaagagatattg    60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21080 atgttcaaaccccaaggattgtatgattacatatgccaacagtggcaagaagagatattg 21139


Query: 61    ccaagtttatgtgactacataaaaatccctaataaatctcctcactttgatgcaaaatgg   120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21140 ccaagtttatgtgactacataaaaatccctaataaatctcctcactttgatgcaaaatgg 21199


Query: 121   gaagaacatggttatatggagcaggcagttaatcacattgccaattggtgtaagtcgcat   180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21200 gaagaacatggttatatggagcaggcagttaatcacattgccaattggtgtaagtcgcat 21259


Query: 181   gctcccaaaggaatgactctggaaattgttcgcctgaaaaataggactccattactattt   240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21260 gctcccaaaggaatgactctggaaattgttcgcctgaaaaataggactccattactattt 21319


Query: 241   atggaaattccaggccaaattgatgacactgtgttgctttatgggcacttggataaacaa   300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21320 atggaaattccaggccaaattgatgacactgtgttgctttatgggcacttggataaacaa 21379


Query: 301   cctgagatgtcaggctggagtgacgatttacatccatggaaacccgtattgaaaaatgga   360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21380 cctgagatgtcaggctggagtgacgatttacatccatggaaacccgtattgaaaaatgga 21439


Query: 361   ttgttatacggaagaggaggggcagatgatggatattctgcttatgcatcactcacggct   420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21440 ttgttatacggaagaggaggggcagatgatggatattctgcttatgcatcactcacggct 21499


Query: 421   attcgcgccttggaacagcaaggtttgccatatcctcgttgtatattaatcatcgaagcg   480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21500 attcgcgccttggaacagcaaggtttgccatatcctcgttgtatattaatcatcgaagcg 21559


Query: 481   tgtgaggaaagtggcagttacgatttgccttttttatattgagttgctgaaagagcgtatt   540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21560 tgtgaggaaagtggcagttacgatttgccttttttatattgagttgctgaaagagcgtatt 21619


Query: 541   ggtaaaccatcattggttatttgtcttgattccggagcaggtaattatgagcagttatgg   600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21620 ggtaaaccatcattggttatttgtcttgattccggagcaggtaattatgagcagttatgg 21679


Query: 601   atgactacgtcattacgcggtaatttggtcggtaagttaactgttgaattaattaatgag   660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21680 atgactacgtcattacgcggtaatttggtcggtaagttaactgttgaattaattaatgag 21739
```

```
Query: 661   ggcgttcattctgggagcgccagtggtatagtggcagacagtttcagagtagctcggcaa   720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21740 ggcgttcattctgggagcgccagtggtatagtggcagacagtttcagagtagctcggcaa 21799

Query: 721   ttgatcagcaggatagaggacgaaaacaccggagagataaaattacctcagttgtattgt   780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21800 ttgatcagcaggatagaggacgaaaacaccggagagataaaattacctcagttgtattgt 21859

Query: 781   gatattcctgatgagagaataaaacaagcgaaacaatgtgcggaaattctaggtgaacaa   840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21860 gatattcctgatgagagaataaaacaagcgaaacaatgtgcggaaattctaggtgaacaa 21919

Query: 841   gtttatagcgaatttccatggatagattctgccaaacccgttattcaagacaaacagcaa   900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21920 gtttatagcgaatttccatggatagattctgccaaacccgttattcaagacaaacagcaa 21979

Query: 901   ttaatattaaacagaacatggcgccctgccttgacggtgactggtgcagatgggtttcca   960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21980 ttaatattaaacagaacatggcgccctgccttgacggtgactggtgcagatgggtttcca 22039

Query: 961   gcgatagctgatgcagggaacgtaatgcgccctgttacgtctttgaaattatccatgcgc  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22040 gcgatagctgatgcagggaacgtaatgcgccctgttacgtctttgaaattatccatgcgc 22099

Query: 1021  cttccaccactggttgatccagaagcagcttctgttgctatggaaaaagccctgacccaa  1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22100 cttccaccactggttgatccagaagcagcttctgttgctatggaaaaagccctgacccaa 22159

Query: 1081  aaccctccctataatgcaaaggttgattttaaaatacaaaatggagggtccaagggatgg  1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22160 aaccctccctataatgcaaaggttgattttaaaatacaaaatggagggtccaagggatgg 22219

Query: 1141  aatgctcctttgctttccgattggttagcgaaagcggcatctgaagcatcaatgacttat  1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22220 aatgctcctttgctttccgattggttagcgaaagcggcatctgaagcatcaatgacttat 22279

Query: 1201  tatgataaacctgctgcttacatgggagaggggggcaccattccatttatgagtatgcta  1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22280 tatgataaacctgctgcttacatgggagaggggggcaccattccatttatgagtatgcta 22339

Query: 1261  ggcgagcaatttcccaaagcacaatttatgataactggtgttttaggcccccattccaat  1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22340 ggcgagcaatttcccaaagcacaatttatgataactggtgttttaggcccccattccaat 22399

Query: 1321  gctcatggtccgaacgagttcttacatttggacatggtaaaaaaactcacctcatgtgtc  1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22400 gctcatggtccgaacgagttcttacatttggacatggtaaaaaaactcacctcatgtgtc 22459

Query: 1381  tcgtacgttctttatagtttttcacagaaaaaataa  1416
             ||||||||||||||||||||||||||||||||||
Sbjct: 22460 tcgtacgttctttatagtttttcacagaaaaaataa 22495
```

```
TBLASTN 2.2.6 [Apr-09-2003]


Query= 1865.3 CONTIG=Contig42 POSCDS1=76674 POSCDS2=77765 SENS=p, Seq
Id : 623
         (367 letters)

Database: contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.done


Sequences producing significant alignments:                          Score    E
                                                                     (bits) Value

LpPhiladelphia_Contig49                                                718    0.0

>LpPhiladelphia_Contig49
          Length = 376826

 Score =  718 bits (1853), Expect = 0.0
 Identities = 366/366 (100%), Positives = 366/366 (100%)
 Frame = +1

Query: 1     GNIMSPSIVFTGGGTAGHVTPNIALIKEFRKEGWNVEYIGSVSGIEKEMIEPLDIPFHGV   60
              GNIMSPSIVFTGGGTAGHVTPNIALIKEFRKEGWNVEYIGSVSGIEKEMIEPLDIPFHGV
Sbjct: 20005 GNIMSPSIVFTGGGTAGHVTPNIALIKEFRKEGWNVEYIGSVSGIEKEMIEPLDIPFHGV 20184

Query: 61    SSGKLRRYFSLKNLLDPFKIVLGIIQSSLLFYKIKPDVVFSKGGFVAFPVVVGAWLNRIP   120
              SSGKLRRYFSLKNLLDPFKIVLGIIQSSLLFYKIKPDVVFSKGGFVAFPVVVGAWLNRIP
Sbjct: 20185 SSGKLRRYFSLKNLLDPFKIVLGIIQSSLLFYKIKPDVVFSKGGFVAFPVVVGAWLNRIP 20364

Query: 121   VVAHESDMSPGLANRLSFPFVNKICLTFDAGKKYFKRQDKIEVTGTPIRQQLLTGNRMKG   180
              VVAHESDMSPGLANRLSFPFVNKICLTFDAGKKYFKRQDKIEVTGTPIRQQLLTGNRMKG
Sbjct: 20365 VVAHESDMSPGLANRLSFPFVNKICLTFDAGKKYFKRQDKIEVTGTPIRQQLLTGNRMKG 20544

Query: 181   LELCGFNSSKPCLLVVGGSLGAGSINSCIRSALKQLTSEFQVIHLCGKGKLDSSLVGVEG   240
              LELCGFNSSKPCLLVVGGSLGAGSINSCIRSALKQLTSEFQVIHLCGKGKLDSSLVGVEG
Sbjct: 20545 LELCGFNSSKPCLLVVGGSLGAGSINSCIRSALKQLTSEFQVIHLCGKGKLDSSLVGVEG 20724

Query: 241   YCQFEYANEELADLFAASSVVISRAGANSLYEILALGKPHILIPISSQVSRGDQIQNARY   300
              YCQFEYANEELADLFAASSVVISRAGANSLYEILALGKPHILIPISSQVSRGDQIQNARY
Sbjct: 20725 YCQFEYANEELADLFAASSVVISRAGANSLYEILALGKPHILIPISSQVSRGDQIQNARY 20904

Query: 301   FQGLGISVVIQDELLKADVLLQAVQDVMRKKDEIDNKIKALKIESATDKIVAIIKEQAHV   360
              FQGLGISVVIQDELLKADVLLQAVQDVMRKKDEIDNKIKALKIESATDKIVAIIKEQAHV
Sbjct: 20905 FQGLGISVVIQDELLKADVLLQAVQDVMRKKDEIDNKIKALKIESATDKIVAIIKEQAHV 21084

Query: 361   QTPRIV 366
              QTPRIV
Sbjct: 21085 QTPRIV 21102


  Database: contigsLpPhiladelphia
    Posted date:  Nov 20, 2003 10:38 AM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
   0.321    0.139    0.399

Gapped
Lambda     K      H
   0.267   0.0410    0.140
```

```
Matrix: BLOSUM62

Query= 1865.3 CONTIG=Contig42 POSCDS1=76674 POSCDS2=77765 SENS=p
         (1092 letters)

Database: /home/Gmp/rusniok/projets/legionella/pourBrevet-
191103/contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.....................................................done

Sequences producing significant alignments:                          Score     E
                                                                     (bits)  Value

LpPhiladelphia_Contig49                                                2165  0.0

>LpPhiladelphia_Contig49
          Length = 376826

 Score = 2165 bits (1092), Expect = 0.0
 Identities = 1092/1092 (100%)
 Strand = Plus / Plus


Query: 1      atgagcccaagtattgtttttaccgggggaggaactgccggacatgtaacgcctaatatc   60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20014 atgagcccaagtattgtttttaccgggggaggaactgccggacatgtaacgcctaatatc 20073


Query: 61     gctttgattaaggaatttcgaaaagaaggctggaatgtagaatatatcggctctgtttcc   120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20074 gctttgattaaggaatttcgaaaagaaggctggaatgtagaatatatcggctctgtttcc 20133


Query: 121    ggaattgaaaaggagatgattgagccgctggacattccttttcatggggtcagtagcggt   180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20134 ggaattgaaaaggagatgattgagccgctggacattccttttcatggggtcagtagcggt 20193


Query: 181    aaattgcgcaggtattttagtttgaagaacttgcttgatcctttcaaaattgttctggga   240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20194 aaattgcgcaggtattttagtttgaagaacttgcttgatcctttcaaaattgttctggga 20253


Query: 241    attattcaatcttctttgctattttataaaatcaaacccgatgtggtttttttcaaaaggt   300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20254 attattcaatcttctttgctattttataaaatcaaacccgatgtggtttttttcaaaaggt 20313


Query: 301    ggctttgtagcctttcctgtggttgtaggcgcctggttaaatcgaattcctgttgtcgct   360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20314 ggctttgtagcctttcctgtggttgtaggcgcctggttaaatcgaattcctgttgtcgct 20373


Query: 361    catgagtctgatatgagcccaggacttgcgaatcgcctatccttttccttttcgtcaataaa   420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20374 catgagtctgatatgagcccaggacttgcgaatcgcctatccttttccttttcgtcaataaa 20433


Query: 421    atatgtcttacttttgatgctggcaaaaaatactttaagcgtcaggataaaatagaagtg   480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20434 atatgtcttacttttgatgctggcaaaaaatactttaagcgtcaggataaaatagaagtg 20493
```

```
Query: 481   acgggtactccaattcgtcaacagctattaactggaaatcgaatgaaaggattggagtta   540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20494 acgggtactccaattcgtcaacagctattaactggaaatcgaatgaaaggattggagtta 20553

Query: 541   tgcggatttaattcctccaaaccttgcctgcttgtagtgggaggaagcttaggggctggt   600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20554 tgcggatttaattcctccaaaccttgcctgcttgtagtgggaggaagcttaggggctggt 20613

Query: 601   tcaattaacagttgtattcgaagcgcattgaaacaattgacatcagaatttcaagtcatt   660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20614 tcaattaacagttgtattcgaagcgcattgaaacaattgacatcagaatttcaagtcatt 20673

Query: 661   catctttgtggcaagggaaaacttgattcttcattggttggtgtggagggatattgccaa   720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20674 catctttgtggcaagggaaaacttgattcttcattggttggtgtggagggatattgccaa 20733

Query: 721   tttgaatacgccaatgaagagttggctgatctgttcgctgcttcttctgtggtgatttct   780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20734 tttgaatacgccaatgaagagttggctgatctgttcgctgcttcttctgtggtgatttct 20793

Query: 781   cgagcaggagctaattctttgtatgaaatattagcattaggaaaaccacatatcttaatt   840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20794 cgagcaggagctaattctttgtatgaaatattagcattaggaaaaccacatatcttaatt 20853

Query: 841   ccaatctcttcacaagtaagcagaggagatcaaattcagaatgcaaggtacttccaggga   900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20854 ccaatctcttcacaagtaagcagaggagatcaaattcagaatgcaaggtacttccaggga 20913

Query: 901   ttgggaataagcgttgtgattcaggacgagttattgaaagctgatgttctattacaggca   960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20914 ttgggaataagcgttgtgattcaggacgagttattgaaagctgatgttctattacaggca 20973

Query: 961   gtacaggacgtaatgcgaaaaaaagatgaaatagataataaaatcaaagcattaaaaatt  1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 20974 gtacaggacgtaatgcgaaaaaaagatgaaatagataataaaatcaaagcattaaaaatt 21033

Query: 1021  gagtctgccactgataagattgtggcaattatcaaggagcaagcacatgttcaaacccca 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 21034 gagtctgccactgataagattgtggcaattatcaaggagcaagcacatgttcaaacccca 21093

Query: 1081  aggattgtatga 1092
              ||||||||||||
Sbjct: 21094 aggattgtatga 21105


------------------------------------------------------------


TBLASTN 2.2.6 [Apr-09-2003]


Query= 2066.5 CONTIG=Contig46 POSCDS1=56766 POSCDS2=57173 SENS=p, Seq
Id : 732
         (150 letters)
```

```
Database: contigsLpPhiladelphia
          51 sequences; 3,410,887 total letters

Searching.done

Sequences producing significant alignments:                          Score    E
                                                                     (bits) Value

LpPhiladelphia_Contig49                                                323   4e-90

>LpPhiladelphia_Contig49
          Length = 376826

 Score =  323 bits (828), Expect = 4e-90
 Identities = 149/149 (100%), Positives = 149/149 (100%)
 Frame = +1

Query: 1      IMYLRLLALSALCFVTSPIWSFTCIYTLVKDNCWTDYDVTVDVIEDSTSKTLLTLTAPKG    60
               IMYLRLLALSALCFVTSPIWSFTCIYTLVKDNCWTDYDVTVDVIEDSTSKTLLTLTAPKG
Sbjct: 89377  IMYLRLLALSALCFVTSPIWSFTCIYTLVKDNCWTDYDVTVDVIEDSTSKTLLTLTAPKG 89556

Query: 61     KSWARGTFNCEAAEGLRYVAQFSPVFWQNDVGKTYPALRNWYLPAKVNPGDLAWTIPVCF   120
               KSWARGTFNCEAAEGLRYVAQFSPVFWQNDVGKTYPALRNWYLPAKVNPGDLAWTIPVCF
Sbjct: 89557  KSWARGTFNCEAAEGLRYVAQFSPVFWQNDVGKTYPALRNWYLPAKVNPGDLAWTIPVCF 89736

Query: 121    PADFAQVPFPPNVAGNCKCNFKNIPDPKL   149
               PADFAQVPFPPNVAGNCKCNFKNIPDPKL
Sbjct: 89737  PADFAQVPFPPNVAGNCKCNFKNIPDPKL 89823


  Database: contigsLpPhiladelphia
    Posted date:  Nov 20, 2003 10:38 AM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
   0.323    0.138    0.473

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62


Query= 2066.5 CONTIG=Contig46 POSCDS1=56766 POSCDS2=57173 SENS=p
         (408 letters)

Database: /home/Gmp/rusniok/projets/legionella/pourBrevet-
191103/contigsLpPhiladelphia
          51 sequences; 3,410,887 total letters

Searching.....................................................done


Sequences producing significant alignments:                          Score    E
                                                                     (bits) Value

LpPhiladelphia_Contig49                                                809  0.0

>LpPhiladelphia_Contig49
          Length = 376826

 Score =  809 bits (408), Expect = 0.0
 Identities = 408/408 (100%)
```

```
  Strand = Plus / Plus

 Query: 1     gtgactagcccaatttggtctttcacatgcatctatactttggttaaagacaattgttgg    60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89419 gtgactagcccaatttggtctttcacatgcatctatactttggttaaagacaattgttgg 89478

 Query: 61    actgattatgatgttactgtcgatgtcattgaagattctacgtcaaaaactttgttgaca   120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89479 actgattatgatgttactgtcgatgtcattgaagattctacgtcaaaaactttgttgaca 89538

 Query: 121   cttaccgctcccaaaggaaaatcatgggctagaggtactttcaattgtgaggctgctgaa   180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89539 cttaccgctcccaaaggaaaatcatgggctagaggtactttcaattgtgaggctgctgaa 89598

 Query: 181   gggttgagatatgtcgctcaattttcgcctgtcttttggcaaaatgatgttggaaaaact   240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89599 gggttgagatatgtcgctcaattttcgcctgtcttttggcaaaatgatgttggaaaaact 89658

 Query: 241   tacccggcattaagaaattggtatttaccagcaaaagtgaatcctggagatttggcctgg   300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89659 tacccggcattaagaaattggtatttaccagcaaaagtgaatcctggagatttggcctgg 89718

 Query: 301   actatcccggtttgttttccggcagattttgctcaagttccctttccacctaatgtagca   360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89719 actatcccggtttgttttccggcagattttgctcaagttccctttccacctaatgtagca 89778

 Query: 361   ggaaactgtaagtgcaacttcaagaacattcctgatcccaagctttaa 408
               ||||||||||||||||||||||||||||||||||||||||||||||
 Sbjct: 89779 ggaaactgtaagtgcaacttcaagaacattcctgatcccaagctttaa 89826

----------------------------------------------------------------


TBLASTN 2.2.6 [Apr-09-2003]


Query= 3159.2 CONTIG=Contig46 POSCDS1=34563 POSCDS2=35318 SENS=p, Seq
Id : 1433
         (265 letters)

Database: contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.done


Sequences producing significant alignments:                          Score    E
                                                                     (bits) Value

LpPhiladelphia_Contig49                                                537   e-154

>LpPhiladelphia_Contig49
         Length = 376826

 Score =  537 bits (1383), Expect = e-154
 Identities = 264/264 (100%), Positives = 264/264 (100%)
 Frame = +1
```

```
Query: 1      CVVESFFLILLFPMWKILYQLASPKNFYNYAGRLIPWLAVSALTTMAIGMVWGLVFAPPD    60
              CVVESFFLILLFPMWKILYQLASPKNFYNYAGRLIPWLAVSALTTMAIGMVWGLVFAPPD
Sbjct: 67177 CVVESFFLILLFPMWKILYQLASPKNFYNYAGRLIPWLAVSALTTMAIGMVWGLVFAPPD 67356

Query: 61     YQQGDAYRIIFVHVPSAFLSMALYAWMGFLAILLLVWRIKMAGLLIHKVAQLGACMAFLA   120
              YQQGDAYRIIFVHVPSAFLSMALYAWMGFLAILLLVWRIKMAGLLIHKVAQLGACMAFLA
Sbjct: 67357 YQQGDAYRIIFVHVPSAFLSMALYAWMGFLAILLLVWRIKMAGLLIHKVAQLGACMAFLA 67536

Query: 121    LITGSIWGKPMWGAWWVWDARLTSELILLLLYLAILATYQAVKNKEDGDKIIAILALVGL   180
              LITGSIWGKPMWGAWWVWDARLTSELILLLLYLAILATYQAVKNKEDGDKIIAILALVGL
Sbjct: 67537 LITGSIWGKPMWGAWWVWDARLTSELILLLLYLAILATYQAVKNKEDGDKIIAILALVGL 67716

Query: 181    IDLPIIHYSVYWWNTLHQGATLSVFAKPKIALSMLYPLLITLLGFFLYSLWIILEKARNE   240
              IDLPIIHYSVYWWNTLHQGATLSVFAKPKIALSMLYPLLITLLGFFLYSLWIILEKARNE
Sbjct: 67717 IDLPIIHYSVYWWNTLHQGATLSVFAKPKIALSMLYPLLITLLGFFLYSLWIILEKARNE 67896

Query: 241    VLFRERKQSWVKIQFEEESDESVF   264
              VLFRERKQSWVKIQFEEESDESVF
Sbjct: 67897 VLFRERKQSWVKIQFEEESDESVF 67968


  Database: contigsLpPhiladelphia
    Posted date:  Nov 20, 2003 10:38 AM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
   0.330    0.143    0.468

Gapped
Lambda     K      H
   0.267   0.0410    0.140

Matrix: BLOSUM62

Query= 3159.2 CONTIG=Contig46 POSCDS1=34563 POSCDS2=35318 SENS=p
         (756 letters)

Database: /home/Gmp/rusniok/projets/legionella/pourBrevet-
191103/contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.....................................................done

Sequences producing significant alignments:                          Score     E
                                                                     (bits)  Value

LpPhiladelphia_Contig49                                                1499  0.0

>LpPhiladelphia_Contig49
          Length = 376826

 Score = 1499 bits (756), Expect = 0.0
 Identities = 756/756 (100%)
 Strand = Plus / Plus


Query: 1      atgtggaagatattgtatcagttggcatcgccaaaaaatttttataactacgcgggacgt    60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67216 atgtggaagatattgtatcagttggcatcgccaaaaaatttttataactacgcgggacgt 67275


Query: 61     ctcattccctggttggcagtcagtgctttgactaccatggccattggtatggtttgggga   120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67276 ctcattccctggttggcagtcagtgctttgactaccatggccattggtatggtttgggga 67335
```

```
Query: 121   ttggtatttgctccaccagattatcagcaaggggatgcataccgaattatttttgttcat  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67336 ttggtatttgctccaccagattatcagcaaggggatgcataccgaattatttttgttcat 67395

Query: 181   gtacccagcgctttttttatcaatggcattgtatgcctggatggggtttctggccatttta  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67396 gtacccagcgctttttttatcaatggcattgtatgcctggatggggtttctggccatttta 67455

Query: 241   ttgttggtgtggcgtatcaaaatggcagggcttttgattcataaggtcgcgcaattaggt  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67456 ttgttggtgtggcgtatcaaaatggcagggcttttgattcataaggtcgcgcaattaggt 67515

Query: 301   gcctgcatggcatttcttgctttaattacagggagcatttggggtaaacccatgtggggt  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67516 gcctgcatggcatttcttgctttaattacagggagcatttggggtaaacccatgtggggt 67575

Query: 361   gcctggtgggtatgggatgcccgcctgacctcagaattaatacttttgttgctctatctg  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67576 gcctggtgggtatgggatgcccgcctgacctcagaattaatacttttgttgctctatctg 67635

Query: 421   gcaattctggctacctatcaagcggtaaaaaataaagaagatggagataaaataatagca  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67636 gcaattctggctacctatcaagcggtaaaaaataaagaagatggagataaaataatagca 67695

Query: 481   attttagctttggtgggtttaattgatttaccaataattcattattcagtttattggtgg  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67696 attttagctttggtgggtttaattgatttaccaataattcattattcagtttattggtgg 67755

Query: 541   aatactttacaccaaggtgcaactttatctgtgtttgccaaacccaaaattgctctcagt  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67756 aatactttacaccaaggtgcaactttatctgtgtttgccaaacccaaaattgctctcagt 67815

Query: 601   atgttgtatccattgttaatcactttgctgggtttttttcttgtattccttatggatcatt  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67816 atgttgtatccattgttaatcactttgctgggtttttttcttgtattccttatggatcatt 67875

Query: 661   ttggaaaaagcacgtaatgaagtcttattcagggagagaaagcaatcatgggttaagatt  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67876 ttggaaaaagcacgtaatgaagtcttattcagggagagaaagcaatcatgggttaagatt 67935

Query: 721   caatttgaggaagagtctgatgaatcagtttttttga   756
             ||||||||||||||||||||||||||||||||||||
Sbjct: 67936 caatttgaggaagagtctgatgaatcagtttttttga 67971

----------------------------------------------------------------

TBLASTN 2.2.6 [Apr-09-2003]

Query= 4774.1 CONTIG=Contig46 POSCDS1=50654 POSCDS2=50950 SENS=m, Seq
Id : 2523
```

```
         (103 letters)

Database: contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.done

                                                                    Score    E
Sequences producing significant alignments:                         (bits) Value

LpPhiladelphia_Contig49                                               205   4e-55

>LpPhiladelphia_Contig49
          Length = 376826

 Score =  205 bits (522), Expect = 4e-55
 Identities = 102/102 (100%), Positives = 102/102 (100%)
 Frame = -1

Query: 1     RRSKMPEIHTLDNPYITILTIFVLACFVGYYVVWKVTPALHTPLMSVTNAISSIIILGAL    60
              RRSKMPEIHTLDNPYITILTIFVLACFVGYYVVWKVTPALHTPLMSVTNAISSIIILGAL
Sbjct: 83615 RRSKMPEIHTLDNPYITILTIFVLACFVGYYVVWKVTPALHTPLMSVTNAISSIIILGAL 83436

Query: 61    IAAGSELIGCITWLGGIAIFITSINIFGGFVVTQRMLRMYKK   102
              IAAGSELIGCITWLGGIAIFITSINIFGGFVVTQRMLRMYKK
Sbjct: 83435 IAAGSELIGCITWLGGIAIFITSINIFGGFVVTQRMLRMYKK 83310


  Database: contigsLpPhiladelphia
    Posted date:  Nov 20, 2003 10:38 AM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
   0.330    0.144    0.442

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62


Query= 4774.1 CONTIG=Contig46 POSCDS1=50654 POSCDS2=50950 SENS=m
         (297 letters)

Database: /home/Gmp/rusniok/projets/legionella/pour Brevet-
191103/contigsLpPhiladelphia
           51 sequences; 3,410,887 total letters

Searching.....................................................done

                                                                    Score     E
Sequences producing significant alignments:                         (bits)  Value

LpPhiladelphia_Contig49                                               589   e-169

>LpPhiladelphia_Contig49
          Length = 376826

 Score =  589 bits (297), Expect = e-169
 Identities = 297/297 (100%)
 Strand = Plus / Minus
```

```
Query: 1     atgcctgaaattcatacacttgataatccttatattacaatattaaccattttcgtactg    60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 83603 atgcctgaaattcatacacttgataatccttatattacaatattaaccattttcgtactg 83544

Query: 61    gcctgttttgtaggttattatgtggtttggaaagtaacaccggctttacatacaccccta   120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 83543 gcctgttttgtaggttattatgtggtttggaaagtaacaccggctttacatacaccccta 83484

Query: 121   atgtcagtaaccaatgccatatccagtattattatacttggtgctttaattgctgcagga   180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 83483 atgtcagtaaccaatgccatatccagtattattatacttggtgctttaattgctgcagga 83424

Query: 181   agtgaattgatcggatgcataacctggttaggtggcatagccatattcattacttcaatt   240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 83423 agtgaattgatcggatgcataacctggttaggtggcatagccatattcattacttcaatt 83364

Query: 241   aatatttttggtggctttgtagtaactcaacgcatgcttcgcatgtataaaaaataa 297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 83363 aatatttttggtggctttgtagtaactcaacgcatgcttcgcatgtataaaaaataa 83307
```

Exemple 5: Other examples of alignment fo sequences

4546.3 (Seq ID 2365)
3009.2 (Seq ID 1331)

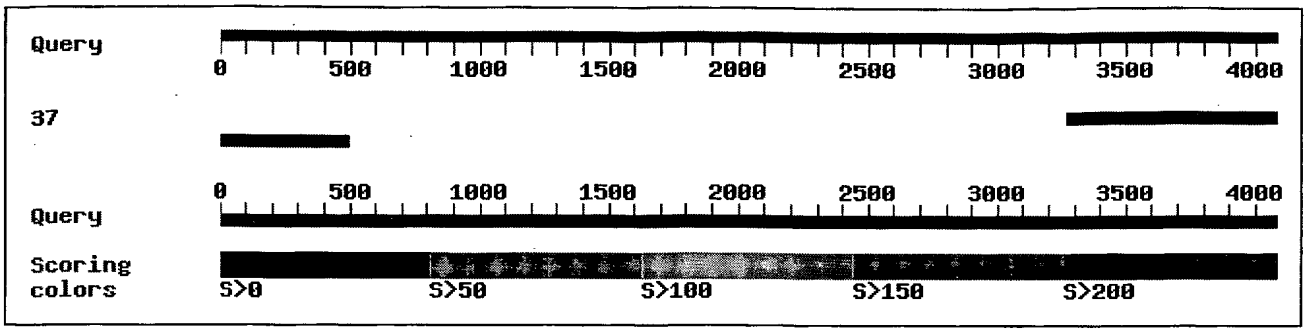
BLASTN 2.2.6 [Apr-09-2003]

Query= Lp Paris Contig48_66441-70516
(4076 letters)

Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
51 sequences; 3,410,887 total letters

Searching.done

Sequences producing significant alignments:	Score (bits)	E Value
37	1439	0.0



>37

Length = 67519

Score = 1439 bits (726), Expect = 0.0
Identities = 784/802 (97%), Gaps = 1/802 (0%)
Strand = Plus / Plus

```
Query: 3275  gcaaatatggtaaaaatgttgacattaattagtggagtatgagatatttatttttgcgagg  3334
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct: 63920 gcaaatatggtaaaaatgttgacattaatcagtggagtatgagatatttatttttgcgagg 63979

Query: 3335  ttggacttgctatgtgtcattccactgggtcaattggcgacacactgattggccctttct  3394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63980 ttggacttgctatgtgtcattccactgggtcaattggcgacacactgattggccctttct 64039

Query: 3395  attatcctgaaatcctgacaagagctctctatggcttaatctataagctgcttgtgatta  3454
              |||||||| ||| |||||||| || ||||||||||||||||||||||||||||||||||
Sbjct: 64040 attatcctaaaaccctgacatgaactctctatggcttaatctataagctgcttgtgatta 64099

Query: 3455  atttcatcgcaatataagccattaaaataccgctaagtaactctatttttgccatactt  3514
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct: 64100 atttcatcgcaatataagccattaaaataccactaagtaactctatttttgccatactt 64159

Query: 3515  tattttgagttaacaggtttgaaaaataacgagtagtcatcgttaatgaactgaaccaaa  3574
              |||||||| |||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64160 tattttgggttaataggtttgaaaaataacgagtagtcatcgttaatgaactgaaccaaa 64219

Query: 3575  tcatgctggcagaaattacccctgctaaaaaagccagtttatggtcaggatattgactgc  3634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64220 tcatgctggcagaaattacccctgctaaaaaagccagtttatggtcaggatattgactgc 64279

Query: 3635  tgccgctacccacaatcaccagggtgtctataatggcgtgaggattaagcagactaaacc  3694
              |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct: 64280 tgccgctgcccacaatcaccagggtatctataatggcgtgaggattaagcagactaaacc 64339

Query: 3695  ccagggcaaataaaatgatttgcattcttgtatgaggctgatgtgtttctacaacggttt  3754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64340 ccagggcaaataaaatgatttgcattcttgtatgaggctgatgtgtttctacaacggttt 64399

Query: 3755  gtttgtttttggacagcgcgcttttaagtttttttattgcataataaattaaaaaggcag  3814
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64400 gtttgtttttggacaacgcgcttttaagtttttttattgcataataaattaaaaaggcag 64459
```

```
Query: 3815  agcctagccataccatccagatttgcaagtttggatgagccaaaagcaattgatgtaaac  3874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct: 64460 agcctagccataccatccagatttgcaagtttggatgagccaaaagcaattgatgtaatc 64519

Query: 3875  cagcgacactgccgcacaccaaaatgacatcacagaaaaagcaaatgacggcagagagag  3934
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct: 64520 cagcgacactgccgcacaccaaaatgacatcgcagaaaaagcaaatgacggcagagagag 64579

Query: 3935  cggcatgatttttttcgcgcaccttgcctgataagaaagacattttgcggacctaaggcca  3994
             | ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct: 64580 ccgcatgatttttttcgcgcaccttgcctgataagaaaaacattttgcggacctaaggcca 64639

Query: 3995  ttatcaaagataatcccaagagtaatccattaaaataaatcaacataattgcatcaggat  4054
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64640 ttatcaatgataatcccaagagtaatccattaaaataaatcaacataattgcatcaggac 64699

Query: 4055  agtaaaaaaaaaggcgattata  4076
             ||| ||||||||||||||||||
Sbjct: 64700 agt-aaaaaaaaggcgattata 64720
```

Score = 957 bits (483), Expect = 0.0
Identities = 495/499 (99%)
Strand = Plus / Plus

```
Query: 1     ctacaaattttgcaaggttattaaatagtggttttcatctggcggcctattgttttttg    60
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct: 63429 ctacaaattttgcaaggtaattaaatagtggttttcatctggcggcctattgttttttg 63488

Query: 61    ggaaagccataagcattctgccaattgatccatgattaaatgttcaacagccatgggatc   120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63489 ggaaagccataagcattctgccaattgatccatgattaaatgttcaacagccatgggatc 63548

Query: 121   ctggtatttttctattaacttggtgtaaacagtacggatgccttgtggcctatcggtcgt   180
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63549 ctgatatttttctattaacttggtgtaaacagtacggatgccttgtggcctatcggtcgc 63608

Query: 181   tacttgttcgcgaacggcaagatgaagccccatatgaaggaagggattggtttcgcctaa   240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63609 tacttgttcgcgaacggcaagatgaagccccatatgaaggaagggattggtttcgcctaa 63668

Query: 241   ttcaggataataagtatgttcaggaaaagattgaatttgttcaatgactttatggtattc   300
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct: 63669 ttcaggataataagtatgttcaggaaaagatggaatttgttcaatgactttatggtattc 63728

Query: 301   cggatgatcaagaatcacttgggcaatttctttttccaagggagaaagttctttttttatt   360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63729 cggatgatcaagaatcacttgggcaatttctttttccaagggagaaagttctttttttatt 63788

Query: 361   ctggtacttattccagctgataaaaaatagctgtcgagtttcttgtactgtatcgccgta   420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63789 ctggtacttattccagctgataaaaaatagctgtcgagtttcttgtactgtatcgccgta 63848
```

```
Query: 421   aaacataatggcccgattgatataaaatgatccattttaactgaataaaaaagtaacaac   480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 63849 aaacataatggcccgattgatataaaatgatccattttaactgaataaaaaagtaacaac 63908


Query: 481   aatgttgatgtgcaaatat   499
              ||||||||||||||||||
Sbjct: 63909 aatgttgatgtgcaaatat 63927


  Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
    Posted date:  Sep 10, 2003 12:44 PM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
    1.37    0.711     1.31

Gapped
Lambda     K      H
    1.37    0.711     1.31


Matrix: blastn matrix:1 -3
```

Alignment of a portion of contig48 (Seq Id 48) of the Paris strain with all the contigs of the Philadelphia strain.

The positions of this fragment in the contig are indicated in the line starting with « Query= ». In this example, the first position (noted 1 in the alignment) of the fragment is thus position 50760 in contig48 of the Paris strain. This fragment terminates in position 56335 in contig48. 50760 should thus be added to the position indicated by the alignment to have the position in the total sequence of the contig. The positions in the contig of the Philadelphia strain are unchanged.

In this exemple, nous pouvons voir que the regions of the Paris strain between the positions 561(+50760) and 2096(+50760) and between 2622(+50760) and 3981(+50760) are absent from the Philadelphia strain. These regions thus contain the following ORFs, specific to the Paris strain:

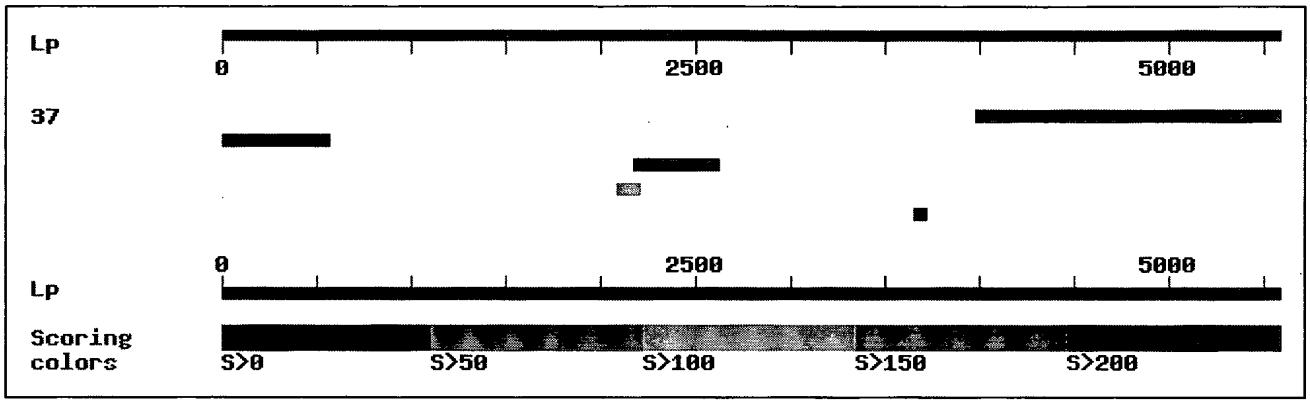
```
322.3 (Seq ID 1466)
321.3 (Seq ID 1462)
3005.2 (Seq ID 1329)
5208.1 (Seq ID 2782)


BLASTN 2.2.6 [Apr-09-2003]


Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
```

programs", Nucleic Acids Res. 25:3389-3402.

Query= Lp Paris Contig48_50760-56335
(5576 letters)

Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
51 sequences; 3,410,887 total letters
Searching.done

Sequences producing significant alignments:	Score (bits)	E Value
37	2892	0.0



>37
Length = 67519

Score = 2892 bits (1459), Expect = 0.0
Identities = 1561/1595 (97%)
Strand = Plus / Plus

```
Query: 3982  cttatggcaaatatttatccctcaaagcgtttctcaataagccattgttaccatgaacca  4041
              ||||||||||||||||||||||||||  ||||||||||||| |||||||||| |||||||
Sbjct: 50163 cttatggcaaatatttatccctcaaaatgtttctcaataagtcattgttaccttgaacca 50222


Query: 4042  gggtaagctttaatttcttaaacaaaatggagtatttagtcctcccttagatggaagct  4101
              |||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||||
Sbjct: 50223 gggtaagctttaatttcttaaacaaaatggagtatatagtcctctcttagatggaagct 50282


Query: 4102  ctttcactgctttgctgagcaaatatttgacgctcctcactgattaattcaatccatttt  4161
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct: 50283 ctttcactgctttgctgagcagatatttgacgctcctcactgattaattcaatccatttt 50342


Query: 4162  ttaggtgtcatgtcgttttgtatcaataagttgatcgctgcactaggatattgactttgt  4221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 50343 ttaggtgtcatgtcgttttgtatcaataagttgatcgctgcactaggatattgactttgt 50402


Query: 4222  gtaatgacaaaataaaagtaagcggtcaatttgcttctgcataaagcccaaccacttttt  4281
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 50403 gtaatggcaaaataaaagtaagcggtcaatttgcttctgcataaagcccaaccacttttt 50462


Query: 4282  gcgacacattgttgcaaactagcaaattcaatttgatcttctttgactaattgttgcagc  4341
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Sbjct: 50463 gcgacacattgttgcaaactaggaaattcaatttgatcttctttgactaattgttgcagc 50522

Query: 4342  ttttgaatcgtgtcagggatacatgcatcaaaaaacaatgtggtcccttgctgtccccat  4401
              ||||||||||||| ||||||| |||||||||||||| ||||||||||||||||||||||
Sbjct: 50523 ttttgaatcgtgttagggatgcatgcatcaaaaaataatgtggtcccttgctgtccccat 50582

Query: 4402  acgtcacttaatataattctttttaaaatgttcatgtcacaaatttgttgctcatgggga  4461
              |||||||||||||||||||||||||||||||| |||||||||||||||| || | |
Sbjct: 50583 acgtcacttaatataattctttttaaaatgttcaagtcacaaatttgttgctcttgcgaa 50642

Query: 4462  ataggtttctcaattctatgtgttaccggttgccccaacatggaatgtgtccattgctga  4521
              |||||||||||||||||||||||||||||||| |||||||||||||||||| |||
Sbjct: 50643 ataggtttctcaattctatgtgttaccggttgccctaacatggaatgtgtccattgttga 50702

Query: 4522  taataagcccgaccgtaaattaatataccacccaaaataaaatcaaaggcagttggggct  4581
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 50703 taataagctcgaccgtaaattaatataccacccaaaataaaatcaaaggcagttggggct 50762

Query: 4582  gcataccaaagagcagtgaagaatttggatttgatgacgtccactttggggtagggatac  4641
              |||||||||||||||||||||| || ||||||||||||||||||||||||||||||||
Sbjct: 50763 gcataccaaagagcagtgaagaaattagatttgatgacgtccactttggggtagggatac 50822

Query: 4642  tgatggttattggttttaatgatatgtttgcgatcgtattgcatttttaaaggttcatat  4701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 50823 tgatggttattggttttaatgatatgtttgcgatcgtattgcatttttaaaggttcatat 50882

Query: 4702  ttaaccaaggatttagcctggtattgattctggccaatacttatcgtttcttccttcttt  4761
              ||||||| |||||||| |||||||||||||| |||||||||||||||||||||||||
Sbjct: 50883 ttaaccagggatttagcttggtattgattctgaccaatacttatcgtttcttccttcttc 50942

Query: 4762  ggtattcctccatcgactattgctaaatgaaagttatgatttaaaacatgttcataaccc  4821
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 50943 ggtactcctccatcgactattgctaaatgaaagttatgatttaaaacatgttcataaccc 51002

Query: 4822  agaaaatgacaatcaggatttttagataaggtatttgctaaatgttgtcgcaatgcctcc  4881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 51003 agaaaatgacaatcaggatttttagataaggtatttgctaaatgttgtcgcaatgcctcc 51062

Query: 4882  ccgactttaatattaggatcataaaaatctggatgtttggctaaattcattacccaaaat  4941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 51063 ccgactttaatattaggatcataaaaatctggatgtttggctaaattcattacccaaaat 51122

Query: 4942  acattgactttattcttttcaccagcaggaatatctcgtaacatcattccccaacaatac  5001
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 51123 acattgactttattcttttcaccagcaggaatatctcgtaacatcattccccaacaatac 51182

Query: 5002  ccaataacttcgtcgtttttatccttggcaataaaacaaatagtctctttataatgtaac  5061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 51183 ccaataacttcgtcgtttttatccttggcaataaaacaaatagtctctttataatgtaac 51242

Query: 5062  attgaagataggacgatatcccctataaaaggcaaatcgccaaaatttgctccatgtgct  5121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
```

```
Sbjct: 51243 attgaagataggacgatatcccctataaaaggcaaatcgccaaaatttgctccatgggct 51302

Query: 5122  aaagaatcgatacgtttaatttcacccattgctttattaagttcagcagtattattggga  5181
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct: 51303 aaagaatcgatacgtttaatttcccccattgctttattaagttcagcagtattattggga 51362

Query: 5182  tctaaaatctgaacttctttgaagccatagggcttgccttccatcatttttaacgaaaaa  5241
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct: 51363 tctaaaatctgaacttctttgaagccatagggcttaccttccatcatttttaacgaaaaa 51422

Query: 5242  ttaaaggtacattcgccatctaacttttgtggtgtttaattttaagaggattttgattt  5301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 51423 ttaaaggtacattcgccatctaacttttgtggtgtttaattttaagaggattttgattt 51482

Query: 5302  ctgcaatttaacagatctgtccataattgatgaatgtaaagtaatctttcatcatttggt  5361
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct: 51483 ctgcaatttaacagatctgtccataattgatgaatgtaaagtaatttttcatcatttggt 51542

Query: 5362  gaggtatttcgcaaattgattaattgctgggtgaaacgatgcacatattctgtatggaga  5421
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct: 51543 gaggtatttcgcaaattgattaattgttgggtgaaacgatgcacatattctgtatggaga 51602

Query: 5422  taaatagcatggctactgaaatcaacaacaaactctgccgtcattaatttttgatttcg  5481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 51603 taaatagcatggctactgaaatcaacaacaaactctgccgtcattaatttttgatttcg 51662

Query: 5482  ggtaagtgatattcaacaggatctgtatatgattttactattgcttctagtgaactccat  5541
              | ||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||
Sbjct: 51663 gataagtgatattcaacaggatctgtatatgattttactattgcctctaatgaactccat 51722

Query: 5542  ttgcgttcaagtacatcgatttttgaatacgacat  5576
              |||||||||||||||||||||||||||||||||
Sbjct: 51723 ttgcgttcaagtacatcgatttttgaatacgacat 51757


 Score =  930 bits (469), Expect = 0.0
 Identities = 543/569 (95%), Gaps = 9/569 (1%)
 Strand = Plus / Plus


Query: 1     atggcgaaatcattgtcgcaattagattctgctaatttgctcccctgttttaacatggaa    60
              ||||||||| |||||||||||| || ||||||||||||||||||||||||||| ||
Sbjct: 46540 atggcgaaaccattgtcgcaattggactctgctaatttgctcccctgttttaacatagag 46599

Query: 61    caagcagaacgcattggaaaacaaatcaataagctcttacagcatgagttttgcgaggaa   120
              |||||||||||||||||||||||||||||| |||||||||||| ||||||||||||
Sbjct: 46600 caagcagaacgcattggaaaacaaatcaataaactcttacagcatgaattttgcgaggaa 46659

Query: 121   aacatcaatccaaagaaatttgcctctatcagtcacaatatcctgcccaaaattatgaca   180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46660 aacatcaatccaaagaaatttgcctctatcagtcacaatatcctgcccaaaattatgaca 46719

Query: 181   gaaacatttttaggagtaaccccgccagaaaactggcagcaattaagcgacgatattata   240
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
```

```
Sbjct: 46720 gaaacatttttaggagtaacccctccagaaaactggcagcaattaagcgacgatattata 46779

Query: 241   aaaaactgcatcgcaaacaagaatctatgcaaaaaagcagctcgcaaagagctggaagaa  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46780 aaaaactgcatcgcaaacaagaatctatgcaaaaaagcagctcgcaaagagttggaagaa 46839

Query: 301   tgcatcaaaccgagaattcctttgatcctgatacaatttggcccctggcttgctcaaaat  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46840 tgcatcaaaccgagaattcctttgatcctgatacaatttggcccctggcttgcccaaaat 46899

Query: 361   tgtccacaattgaataagtctttgattgaacaatggccaaataaacaggctactctcaaa  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46900 tgtccacaattgaataagtctttgattgaacaatggccaaataaacaggccactctcaaa 46959

Query: 421   aagataattaatgaaaacaaaagtgccgagtaatcgagacaggcttaattacgagagtta  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46960 aagataattaatgaaaacaaaagtgctgagtaatcgaggcaggcttaattacgagagtta 47019

Query: 481   tgcctgataaaaccacatttatttacctaat---------ttcatcaaatataactcacc  531
              |||| |||||||| |||||||||||||||||         ||||||||||||||| |||
Sbjct: 47020 tgccggataaaacaacatttatttacctaattcacctaacttcatcaaatataacttacc 47079

Query: 532   gatatgatctacaaccaagttcttaaaac  560
              ||||||||||||||||| ||||||||||
Sbjct: 47080 gatatgatctacaaccaggttcttaaaac 47108


 Score =  656 bits (331), Expect = 0.0
 Identities = 423/451 (93%), Gaps = 2/451 (0%)
 Strand = Plus / Plus


Query: 2172  tatatggccgataaaatttgccagggtcaataaatagtattctgatggttaggtaataaa  2231
              ||||||||||||||||||||||||||||||||||||||||||||| ||| |||||||
Sbjct: 47927 tatatgaccgataaaatttgccagggtcagtaaatagtattctgatggctagataataaa 47986

Query: 2232  tgatgaagagttcttttggtgaaatgaataaaagacccctttttttattgagcgactctta  2291
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 47987 tgatgaaaagttcttttgatgaaatgaataaaagaacc-ttttttattgagcgactctta 48045

Query: 2292  aaaa-gccatttgctttattctgtgcttttggcaagtgacatgatcgcattcaggttttg  2350
              |||| |||||||||||||| ||||||||||||||||    ||| |||||||||| ||||
Sbjct: 48046 aaaaagccatttgctttatcctgtgcttttggcaagcagtatggtcgcattcaggctttg 48105

Query: 2351  tttacatgctaggtgtggtttttttcccgaggcggttgtgagtttgataccataggtttta  2410
              |||||||||||||||| ||||||||| |||||||||||| | ||||||||||||||||||
Sbjct: 48106 tttacatgctaggtgtagtttttttcctgaggcggttgtgggcttgataccataggtttta 48165

Query: 2411  ttgaataggcgccgcttgggaaaaaacaattgtcagtaaggattgcccctgtaatcagat  2470
              |||||||| |||| ||||||||||||||||||| | ||||||||||||||||||| |
Sbjct: 48166 ttgaatagacgccacttgggaaaaaacaattgtcggcaaggattgcccctgtaatcagct 48225

Query: 2471  gggtaagccgattggcttcggcacatctcgaatcagtgacaacctctttcgctttttcat  2530
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
```

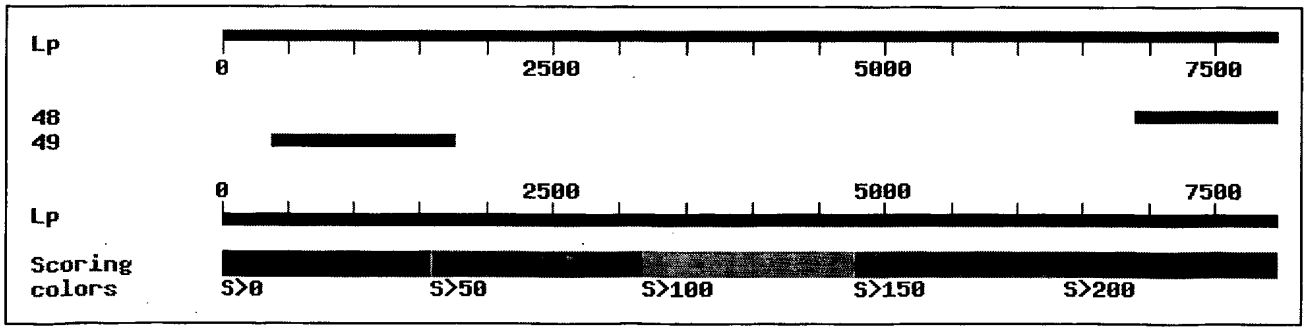
```
Sbjct: 48226 gggtaagccgattggcttcggcacatctcgaatcagtggcaacctctttcgctttttcat 48285


Query: 2531  ctttatttcgcataacaatcctgtgaagttaatctttgcagaggacaccatgatggtttc  2590
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct: 48286 ctttatttcgcataacaatcctgtgaagttaatctttgcagaggacactatgatggtttc 48345


Query: 2591  atgtcatacaaacgaagcaaccggataccga  2621
              |||||||||||||||||||||||| ||||||
Sbjct: 48346 atgtcatacaaacgaagcaaccgggtaccga 48376


 Score =  147 bits (74), Expect = 9e-35
 Identities = 98/106 (92%)
 Strand = Plus / Plus


Query: 2097  acctttctggagtttcccaactacaagatgatactgcgttataataactccatttattat  2156
              ||||| |||| |||||||||||||||||||| || ||||||||||||||||||||||||
Sbjct: 44922 accttcctggcgtttcccaactacaagatgatcctacgttataataactccatttattat 44981


Query: 2157  actggggctatcgagtatatggccgataaaatttgccagggtcaat  2202
               | |||||||||| ||||||||| ||||||||||||||||||||
Sbjct: 44982 gccggggctatcgggtatatggctgataaaatttgccagggtcaat 45027


 Score =  109 bits (55), Expect = 2e-23
 Identities = 64/67 (95%)
 Strand = Plus / Plus


Query: 3653  cgaatccttacaggaaaacgaagcttatggaagtccaacaaggaagaggtagtaaattca  3712
              ||||||||||||||||||||||||||||||||| |||||||||| ||| ||||||||||
Sbjct: 48378 cgaatccttacaggaaaacgaagcttatggaaggccaacaaggaacagggagtaaattca 48437


Query: 3713  taacgcc  3719
              |||||||
Sbjct: 48438 taacgcc 48444


  Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
    Posted date:  Sep 10, 2003 12:44 PM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
    1.37    0.711     1.31

Gapped
Lambda     K      H
    1.37    0.711     1.31


Matrix: blastn matrix:1 -3
```

Alignment of a portion of the contig45 (Seq Id 45) of the Paris strain with all the contigs of the Philadelphia strain.

The positions of this fragment in the contig are indicated in the line beginning with « Query= ». In this example, the first position (noted 1 in the alignment) of the fragment is thus position 44722 in the contig45 of the Paris strain. This fragment terminates at position 52680 in the contig45. It is thus required to add 44722 à to the indicated by the alignment to have the position in the total sequence of the contig. The positions in the contig of the Philadelphia strain unchanged.

In this example, we can see that the region of the Paris strain between the positions 1333(+44722) and 6899(+44722) is absent from the Philadelphia strain. This region thus contains the following ORFs, specific to the Paris strain:

```
4927.1 (Seq ID 2623)
413.5 (Seq ID 2069)
415.2 (Seq ID 2087)
417.3 (Seq ID 2102)


BLASTN 2.2.6 [Apr-09-2003]


Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

Query= Lp Paris Contig45_44722-52680
         (7959 letters)

Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
           51 sequences; 3,410,887 total letters

Searching.done

                                                                  Score    E
Sequences producing significant alignments:                       (bits) Value

48                                                                  1943   0.0
49                                                                    94   2e-18
```




```
>48
          Length = 263853
```

Score = 1943 bits (980), Expect = 0.0
Identities = 1040/1060 (98%)
Strand = Plus / Minus

```
Query: 6900  attctaatggtggccagaggcagaatcgaactgccgacacgaggattttcagtcctctgc  6959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 55098 attctaatggtggccagaggcagaatcgaactgccgacacgaggattttcagtcctctgc 55039

Query: 6960  tctaccgactgagctatctggccactcaaggcttgctattaaactccgagatcgactttg  7019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 55038 tctaccgactgagctatctggccactcaaggcttgctattaaactccgagatcgactttg  54979

Query: 7020  agtcaagtgttgattaagtttttttttgaaaataatttttttttggtaggctatggcaatct  7079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct: 54978 agtcaagtgttgattaagtttttttttgaaaataatttttttttggtaggctatggcaattt 54919

Query: 7080  ctgctcctagaaaagcgttgtttttgataagctcaatattagcttccaggctcttgcctg  7139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Sbjct: 54918 ctgctcctagaaaagcgttgtttttgataagctcaatattagcttccaggctcttgcccg 54859

Query: 7140  cagttaattcagcaatacgttttaacaggaaggggttaatgattttccgctcatgtgtt  7199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 54858 cagttaattcagcaatacgttttaacaggaaggggttaatgattttccgctcatgtgtt 54799

Query: 7200  tggcttcttcatgagcctgttttatgtatggactgatttcctcatcagatagttccgctg  7259
              ||||||| |||||||||||||||||||||||||||||||||||| |||||||| ||||
Sbjct: 54798 tggcttcatcatgagcctgttttatgtatggactgatttcctcatcggatagttcggctg 54739

Query: 7260  atactggaatggggtttgcgacgactattccgtttttcatgttcaatttctgttgaattg  7319
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||
Sbjct: 54738 atactggaatggggtttgcgacgactattccatttttcatattcaatttctgttgaattg 54679

Query: 7320  acatgagatttgctacttcctcgaccgaatttaagcgttgtgggactggtattccactcg  7379
              |||||||||||||||||||||| |||| ||||||||||||| ||||||||||||||||
Sbjct: 54678 acatgagatttgctacttcctcggccgagtttaagcgttgtggaactggtattccactcg 54619

Query: 7380  atctgctgtaaaaagcagggaattcgtctgtggcataacctatgaccggcaccccaaacg  7439
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 54618 atctgctataaaaagcagggaattcgtctgtggcataacctatgaccggcaccccaaacg 54559

Query: 7440  tttcaagaacttccaatgtttttggtaagtcgagaatcgattttgcgccagaacagacta  7499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 54558 tttcaagaacttccaatgtttttggtaagtcgagaatcgattttgcgccagaacagacta 54499

Query: 7500  cggtaactggcgtattggatagttctataagatcagctgaaatatcaaaactcattgtca  7559
              |||||||||||||||||||| ||||||||||||||||||||||||||||||| ||||
Sbjct: 54498 cggtaactggcgtattggatagctctataagatcagctgaaatatcaaaactcatcgtca 54439

Query: 7560  cgtcttgatgaacaccacctatacctccggtgacaaataggggagcctagccatatggg  7619
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct: 54438 cgtcttgatgaacaccacctatacctccggtgacaaataggggagcttagccatatggg 54379
```

```
Query: 7620  cacagaacatggttgctgctaccgttgtgctggcagtcactttacgagataatacaaaag  7679
              ||||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct: 54378 cacaaaacatggttgctgctaccgtcgtgctggcagtcactttacgagataatacaaaag 54319

Query: 7680  aaatgtctctgcgagaggcttttattacttctttttgcagcgcgagatgctccataactt  7739
              ||||||||||||||||||||||||||||||||||||| |||||||||| ||||||||||
Sbjct: 54318 aaatgtctctgcgagaggcttttattacttctttttgaagcgcgagatgttccataactt 54259

Query: 7740  cttgagttaaaccaatacggattttcccttggtgcatcgctatagtagctggaatggcgc  7799
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 54258 cttgagttaaaccaacacggattttcccttggtgcatcgctatagtagctggaatggcgc 54199

Query: 7800  cttgtctacgaataatattttcaacttctattgccgtagttaaattatcagggtagggca  7859
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct: 54198 cttgtctacgaataatattttcaacctctattgccgtagttaaattatcagggtagggca 54139

Query: 7860  ttccatgagagataatggtcgactcaagagcaacaattgggtttttatcattgatagcat  7919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 54138 ttccatgagagataatggtcgactcaagagcaacaattgggtttttatcattgatagcat 54079

Query: 7920  ccagtacttcttcgttaaattccaacaagtcatgaaacat  7959
              |||||||||||||||||||||||||||||||||||||
Sbjct: 54078 ccagtacttcttcgttaaattccaacaagtcatgaaacat 54039
```

>49

Length = 376826

Score = 93.7 bits (47), Expect = 2e-18
Identities = 102/119 (85%), Gaps = 1/119 (0%)
Strand = Plus / Plus

```
Query: 1215   atttgccctgtgtattgtttagtgttggtcgagcggttcactctctgttgaaacccggta  1274
               |||||||||||  | | ||||||||||  ||||||| ||||||   | |||||||  ||
Sbjct: 207683 atttgccctgtctttggtttagtgttgaacgagcggctcactcaaagatgaaacctggcc 207742

Query: 1275   aaa-ccgtaaagctcgaagaaggggggcaaatcaatcgttataaggcaaacgatcccgcc   1332
               ||| || |||||||||||| ||||||||||||||||||||||||||||||||||| ||||
Sbjct: 207743 aaagccataaagctcgaagtagggggcaaatcaatcgttataaggcaaacgatctcgcc  207801
```

Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
Posted date: Sep 10, 2003 12:44 PM
Number of letters in database: 3,410,887
Number of sequences in database: 51

Lambda	K	H
1.37	0.711	1.31

Gapped

Lambda	K	H
1.37	0.711	1.31

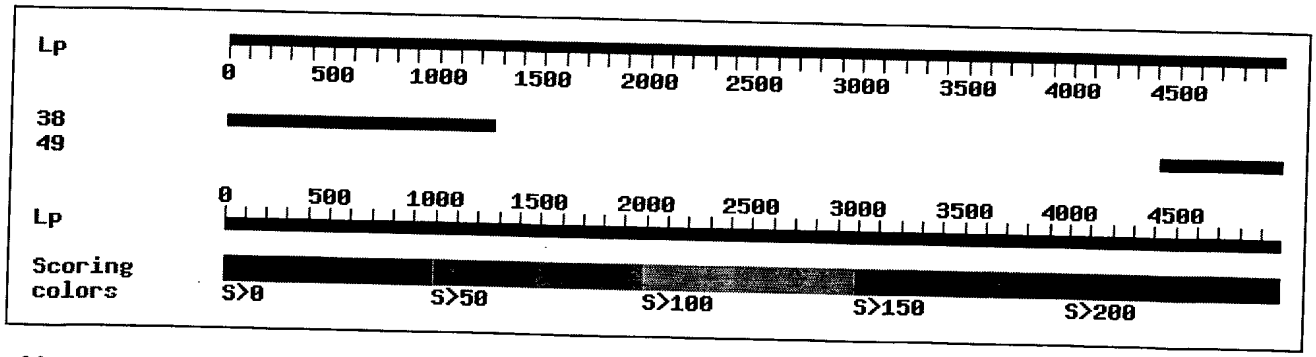
Matrix: blastn matrix:1 -3

Alignment of a portion of contig39 (Seq Id 39) of the Paris strain with all the contigs of the Philadelphia strain.

The positions of this fragment in the contig are indicated in the line starting with « Query= ». In this example, the first position (noted 1 in the alignment) of the fragment is thus position 3990 in contig39 of the Paris strain. This fragment terminates at position 8972 in contig39. 3990 should this be added to the position indicated by the alignment to have the position in the total sequence of the contig. The positions in the contig of the Philadelphia strain are unchanged.

In this example, we can see that the region of the Paris strain between the positions 1264(+3990) and 4465(+3990) is absent from the Philadelphia strain. This region thus contains the following ORFs, specific to the Paris strain:

```
3396.1 (Seq ID 1588)
3395.2 (Seq ID 1587)
3394.1 (Seq ID 1586)


BLASTN 2.2.6 [Apr-09-2003]


Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs",  Nucleic Acids Res. 25:3389-3402.

Query= Lp Paris Contig39_3990-8972
         (4983 letters)

Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
           51 sequences; 3,410,887 total letters

Searching.done


Sequences producing significant alignments:                    Score    E
                                                               (bits) Value

38
49                                                               2123   0.0
                                                                  638   0.0
```




```
>38
          Length = 17003
```

```
Score = 2123 bits (1071), Expect = 0.0
Identities = 1215/1263 (96%)
Strand = Plus / Minus

Query: 1     ctttggttctacatgagcttgcctgagatgtttgcccttatcgattgcaacaatttttac    60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 12277 ctttggttctacatgagcttgcctgagatgtttgcccttatcgattgcaacaatttttac 12218

Query: 61    gccagttgtgagcgtttgtttcgtcctgatttaaaggacgtccccatcgtggtgctatcc   120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 12217 gccagttgtgagcgtttgtttcgtcctgatttaaaggacgtccccatcgtggtgctatcc 12158

Query: 121   aataatgatggctgttgtatcgcacgctcgaatgaagccaaagcattgggcattgccatg   180
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 12157 aataacgatggctgttgtatcgcacgctcgaatgaagccaaagcattgggcattgccatg 12098

Query: 181   ggcgagccgtacttcaaaattaaacatttgtgcaaacagcatggagtgaaagctttttcc   240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 12097 ggcgagccgtacttcaaaattaaacatttgtgcaaacagcatggagtgaaagctttttcc 12038

Query: 241   tcaaattatacgctgtatggcaacatgagtcatcgtgtgatgtgcactattgaagaagcc   300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 12037 tcaaattatacgctgtatggcaacatgagtcatcgtgtgatgtgcactattgaagaagcc 11978

Query: 301   tggccccatatagaagtttactcgattgatgaagcgtttcttgatttaaggagtttaccg   360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 11977 tggccccatatagaagtttactcgattgatgaagcgtttcttgatttaaggagtttaccg 11918

Query: 361   gttgatagccatgattcgttttgcgagcagttacaaaagaaaatcttgaagcacacagga   420
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 11917 gttgatagctatgattcgttttgcgagcagttacaaaagaaaatcttgaagcacacagga 11858

Query: 421   atacccacttccatcggtattggacctactaaaacactagctaaagccgccaatcattta   480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 11857 atacccacttccatcggtattggacctactaaaacactagctaaagccgccaatcattta 11798

Query: 481   tgcaaaaaagtttataaaatccctgtgtttaatatcacctcgaatcgtgggcggttattg   540
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct: 11797 tgcaaaaaagtttataaaattcctgtgtttaatatcacctcgaatcgtgggcggttattg 11738

Query: 541   caacagatttccgttggggacatttggggagtagggcggcaatgggccaataaattaatt   600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 11737 caacagatttccgttggggacatttggggagtagggcggcaatgggccaataaattaatt 11678

Query: 601   tcgcgaggcattcatacggcttatgatttggcaatgaccaatcctcaccttctgaagaaa   660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 11677 tcgcgaggcattcatacggcttatgatttggcaatgaccaatcctcaccttctgaagaaa 11618

Query: 661   tgttttaacgtcgtgttgatgcgtaccgccatggagcttcaaggaattgcttgtggcggt   720
              |||||||||| |||||||||||| || || ||||||||||||||||||||||||||||||
Sbjct: 11617 tgttttaacgccgtgttgatgcgaactgctatggagcttcaaggaattgcttgtggcggt 11558
```

```
Query: 721   ttagaggcaatagagcctaagcaaagtattatgtcatctaaaagttttggtcagatgcaa   780
              ||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct: 11557 ttagaggtaatagagcctaagcaaagtattatgtcatccaaaagttttggtcagatgcaa 11498

Query: 781   actcaacttgcttcgattgaggaatcaatcagtagccattgtgcccgtgcggtggagaaa   840
              |||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct: 11497 actcaaattgcttcgattgaggaatcaatcagtagccattgtgctcgtgcggtggagaaa 11438

Query: 841   atgcgtcgccagcaattggtggcgaagcgcctggttgtatttgtgcatacgaaccgattt   900
              |||||||||||| |||||||||||||| ||| ||||||||||||||||||||||||||||
Sbjct: 11437 atgcgtcgccaacaattggtggcgacgcgtctggttgtatttgtgcatacgaaccgattt 11378

Query: 901   cgcgaagatttggcacagcactttcagtccatcgaatttaagctgattaatcctacagat   960
              || |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 11377 cgggaagatttggctcagcactttcagtccatcgaatttaagctgattaatcctacagat 11318

Query: 961   gatttgcgcttaattaccaaaatggccaagcgatgtctgcaacgcatttttaaaccaggg   1020
              |||||||||||||||||||||||| |||||  ||||||||||||||||||||||||||||
Sbjct: 11317 gatttgcgcttaattaccaaaatagccaaaagatgtctgcaacgcatttttaaaccaggg 11258

Query: 1021  tattactataaaaaggcaggagtatgtcttgaggacttaattcccaaaaacccacgacag   1080
              |||||||||||||||||||| |||||||||| |||||||||| ||||| |||||||||
Sbjct: 11257 tattactataaaaaggcaggggtatgtcttgaagacttaattcctaaaaaaccacgacag 11198

Query: 1081  ctggatatgttttatcaaccaagtgacgagcatctaaaccacacggaacaattgatggcg   1140
              |||||||||||| |||||||||||| |||||||||| ||||| ||||||||||||||
Sbjct: 11197 ctggatatgtttcatcaaccaagtgatgagcatctaaaacacaccgaacaattgatgggt 11138

Query: 1141  gtctttgaccaaatcaatcaaaaatacggacgaagtacaatccgcctcgcggcagagggt   1200
              |||||||||||||||||||||| || ||| |||| || || ||  | || || || ||
Sbjct: 11137 gtctttgaccaaatcaatcaaaagtatggaagaagcacgattcggttagccgccgaaggc 11078

Query: 1201  tattcaaaaccttgggcgatgcgtgctgaactgaaatcgcctgcctataccacacgatgg   1260
              |||||||||||| |||| |||||||||||| |||||||| ||||| |||||||| ||||||
Sbjct: 11077 tattcaaaaccctgggagatgcgtgctgagctgaaatcacctgcttataccacgcgatgg 11018

Query: 1261  tct 1263
              |||
Sbjct: 11017 tct 11015


>49
          Length = 376826

 Score =  638 bits (322), Expect = 0.0
 Identities = 469/518 (90%)
 Strand = Plus / Plus


Query: 4466   gaacaataatcactgataaaaagatcttgagcaaaagtctcaaaatcaaaatagcagatc  4525
               |||||||||||||| |||||||||| |||||||| |||||||||||||||||||| ||
Sbjct: 203775 gaacaataatcactaataaaaagatcccgagcaaaagcctcaaaatcaaaatagcagctc 203834


Query: 4526   aaactatccggcatttgatgggcataacagtcatcaaataattgcctggcaaaatcaacc 4585
               |||||||||||| |||||||||||||||||||||||||| || ||||||||||||||||
```

```
Sbjct: 203835 aaactatccggtatttgatgggcataacagtcatcaaataactgtctggcaaaatcaacc 203894

Query: 4586   tctgaatcataacaaccttggtaatgatcctccaacatggtttgtgcatcatctaccgag 4645
              |||||||||||| ||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct: 203895 tctgaatcatagcaaccttggtaatgatcctccaacatggtttgtgcatcttctaccgaa 203954

Query: 4646   taatcacagagcagggctaatcctaactctccgtgttcttgaatgaatgaagcatattcc 4705
              ||||| ||||| |||||||| |||| |||||| |||||||||| |||||||||| |||
Sbjct: 203955 taatcgcagagaagggctaagcctagctctccatgttcttgaataaatgaagcatactcc 204014

Query: 4706   acaatgttgcttattccttcatattcatgaattctcatgctgccgaaaccttcataatcg 4765
              ||||| || ||||||   || || ||||| ||| | |||||||||| ||||||||||
Sbjct: 204015 acaatattacttattgactcgtactcatggattttgatgctgccgaatccttcataatcg 204074

Query: 4766   tggatggcaaattcctcagcattggtttctgggctattatccaacatttcccagatttct 4825
              || |||||| |||| |||||||||  || ||||| ||||||| ||||||||||||||||
Sbjct: 204075 tgtatggcatattcttcagcattgggctcagggctgttatccagcatttcccagatttct 204134

Query: 4826   ttcatgatgtcatcttcgctttgagtagcatctatccatacaccatgcagtatggcgttg 4885
              | ||| ||||||||||||||||| || |||||||||| |||||||||| ||||| |||
Sbjct: 204135 tccataatgtcatcttcgctttgtgtggcatctatccaaacaccatgcaggatggcattg 204194

Query: 4886   ttgtaagaggctaaacaggcgacgtagattgaaggggtgtccatgggattatctccttgt 4945
              ||||| |||||||||| ||||||||||||||||||||||||||||||||| |||||||
Sbjct: 204195 ttgtatgaggctaaacacgcgacgtagattgaaggggtgtccatgggattatttccttgt 204254

Query: 4946   attaagggagctatcccacacgggagcttgctcccgtg 4983
              |||||||||||| |||||||||||||||| ||||||||
Sbjct: 204255 attaagggagcaatcccacacgggagcttactcccgtg 204292


  Database: /local/http/htdocs/IPF_Gmp/legionella/blastdb/contig
    Posted date:  Sep 10, 2003 12:44 PM
  Number of letters in database: 3,410,887
  Number of sequences in database:  51

Lambda     K      H
    1.37    0.711     1.31

Gapped
Lambda     K      H
    1.37    0.711     1.31


Matrix: blastn matrix:1 -3
```

Example 6: RESULTS

See hereinbelow Tables V, VI, I and II hereinbelow.

Example 7 : Sequencing and annotation of the genome of L. pneumophila Lens strain

Comparison of the sequences of the genomes of *L. pneumophila* Paris strain, Lens strain and Philadelphia strain (http://genome3.cpmc.columbia.edu/~legion/index.html), three strains of serogroup 1, shows that around 88% of these genomes are very strongly preserved (95 to 100% of proteic identity), whereas the remaining 12% are specific to each strain. These results suggest that there is a large genomic diversity at the very centre of the *L. pneumophila* species.

The Table XVI hereinbelow comprises for each of the ORFs identified in the Lens strain its position on the genome, the existence of a peptide signal, the best result of the blast on nrprot (Best-Blastp). The ORFs specific to the strain *L. pneumophila* Lens relative to the strain *L. pneumophila* Paris were identified in considering as specific the ORFs having a percentage of proteic homology less than 75%. In the case where the ORF is preserved in the two genomes, the percentage of homology between the two proteins is mentioned. Finally, the ORFs specific to the *Legionella* genre were identified in considering as specific the ORFs having a percentage of proteic homology with sequences of the bank nrprot less than 25 %.

In conclusion, these results help define DNA probes for developing a typing tool. The utilization of this tool on a large number of strains isolated from patients and strains isolated from the environment can confirm if this tool can predict the risk associated with a strain by definitely discriminating the strains isolated from patients of the other strains.

Material supplied:

- The complete sequence of the genome of *Legionella pneumophila* Lens strain made up of the long chromosome of around 3.33 Mb and of a long plasmid of around 60 kb.
- A list of specific coding phases of *L. pneumophila* Lens strain annotated with their nucleotidic sequences.

Materials and methods.

1. Construction of the shotgun bank of small fragments (size 1.5 to 2.5 kb)

The chromosomal DNA of the strains studied was prepared by a classic method including proteinase K treatment and phenol extraction (9). Around 40 ug of DNA were broken by nebulization (1 minute under pressure of 1 bar) (4). The ends of the fragments of DNA were rendered free by having the DNA-polymerase of the bacteriophage T4 act for 15 minutes at 37°C in the presence of the 4 tri-phosphate nucleotides. The enzyme was inactivated by incubation of 15 mn at 75°C. Adaptors (invitrogen Cat. N° 408-18) have then been ligatured to these ends. After ligature, the fragments of chromosomal DNA having a size between 1500 and 2500 base pairs were purified after electrophoresis on agarose gel. The vector utilized for construction of the bank, pcDNA2.1 (Invitrogen), was digested by the enzyme BstX1 and purified by geneclean (BIO-101) after electrophoresis on agarose gel. The chromosomal DNA and the purified vector were ligatured by action of the ligase of the bacteriophage T4. The ligation mixture was introduced by transformation in the strain of Escherichia coli XL2-blue (Stratagene). About 4000 colonies are obtained per ul of the ligation mixture.

2. Preparation of plasmids and sequencing

The plasmids were prepared from bacterial colonies with the «TempliPhi DNA sequencing template amplification» kit marketed by Amersham Bioscience. The chromosomal inserts were sequenced from their two ends by utilizing the universal primer T7 by following the recommendations of the supplier (Applied-Biosystems). The sequences were determined by utilizing automatic sequencers of type 3700 (Applied-Biosystem).

3. Assembling of sequences

The sequences were assembled by utilizing the set of software developed at the University of Washington, Phred, Phrap and Consed (5, 8). The sequence was finished by utilizing the set of CAAT-box software (7). The finishing stage corresponds to resequencing the regions where the sequence is only slightly secure and sequencing of the regions situated between the contigs. It was done either by sequencing PCR products or by operating on the clones of the bank. The oligonucleotidic sequences were defined by utilizing consed and Primo software (8, 10).

3. Annotation of sequences

The identification of the coding phases (CDS) was completed by utilizing the set of CAAT-box software (7). This program combines the results of different methods: (i)

identification of open reading phases and their tri as a function of their size, (ii) analysis of the probability of being coded by utilizing Genemark software (11), (iii) identification of a start in translation (initiation codon and fixing sequence of the ribosome), (iv) similarity of the proteic sequence deduced with the proteic sequences contained in the sequence banks by utilizing BLASTP software.

The functions of the proteins coded by the coding phases identified were predicted by analysis of the research results of similarities in the non-redundant bank of the NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) by utilizing BLASTP software (1).

4. Comparison of the genomes – identification of the CDS specific to the strain of *L. pneumophila* Paris strain

The set of proteic sequences deduced from the predicted coding phases each genome was compared to the set of proteic sequences possibly coded by the other genome by utilizing BLASTP software. A threshold of 75% of identity on the totality of the length of the protein was retained for identifying the proteins specific to an isolate. This very high value was kept since it best allows the orthologous genes of the paralogous genes (6) to be discriminated. For the proteic sequences for which the sequence preservation is high (> at 70%) the preservation of the nucleotidic sequences of the genes will also be high and could give a signal in hybridization conditions of low stringency. It will be necessary to take into consideration this eventuality in the analysis of the test result.

5. Examples of annotations

5.1. Genes specific to *L. pneumophila* Lens strain. There is no significant similarity between the nucleotidic sequence of the gene of *L. pneumophila* Lens strain and the genome of *L. pneumophila* Paris strain.

ID of *L. pneumophila* Lens strain gene	ID of *L. pneumophila* Lens strain gene (best score)	% of identity of proteic sequences	% of identity of nucleotidic sequences
2795.1	-	-	-
560.1	-	-	-
116.1	-	-	-
3866.1	2661.2	26%	-
2141.1	152.3	24%	not significant

5.2. Genes common to the two strains for which the similarity (identity) of the deduced proteic sequences is less than 75% and value of the similarity at the nucleotide level.

ID of *L. pneumophila* Lens strain gene	ID of *L. pneumophila* Paris strain gene (best score)	% of identity of proteic sequences	% of identity of nucleotidic sequences
2518.1	5987.2	42%	32%
3820.1	3661.4	42%	15%

5.4. Genes common to L. pneumophila Lens strain and Paris strain.

ID of *L. pneumophila* Lens gene	ID of *L. pneumophila* Paris gene (best score)	% of identity of proteic sequences	% of identity of nucleotidic sequences
795.1	3838.3	99%	98%
2457.1	3282.3	100%	98%

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. [Review] [90 refs]. *Nucleic Acids Research.* 25:3389-402.

2. Aurell, H., J. Etienne, F. Forey, M. Reyrolle, P. Girardo, P. Farge, B. Decludt, C. Campese, F. Vandenesch and S. Jarraud. 2003. *Legionella pneumophila* serogroup 1 strain Paris: Endemic distribution throughout France. J Clin Microbiol. 41: 3320-3322

3. Birnboim, H. C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.* 100:243-255.

4. Buchrieser, C., C. Rusniok, L. Frangeul, E. Couve, A. Billault, F. Kunst, E. Carniel, and P. Glaser. 1999. The 102 kb locus of *Yersinia pestis* : sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect. Immun.* **67**:4851-4861.

5. Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8:186-194.

6. Fitch, W. S. 1970. Distingishing homologous from analogous proteins. *Syst. Zool.* 19:99-113.

7. Frangeul L, Glaser P, Rusniok C, Buchrieser C, Duchaud E, Dehoux P and Kunst F. 2004. CAAT-Box, Contigs-Assembly and Annotation tool-box for genome sequencing projects. *Bioinformatics*. in press.

8. Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8:195-202.

10. Li, P., K. C. Kupfer, C. J. Davies, D. Burbee, G. A. Evans, and H. R. Garner. 1997. PRIMO: A primer design program that applies base quality statistics for automated large-scale DNA sequencing. *Genomics*. 40:476-485.

11. Lukashin, A. V., and M. Borodovsky. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 15:1107-1115.

12. Xu, S.Y., and Fomenkov, A. 1994. Construction of pSC101 derivatives with Camr and Tetr for selection or LacZ for blue/white screening. *Biotechniques* **17:**57.

Example 8: Proof in the genome of Legionella pneumophila of the exploitation of functions of the host cell and of the high genomic plasticity

Legionella pneumophila, the causal agent of legionnaires disease, replicates in the form of an intracellular parasite of the amoeba and persists in the environment in the form of a free living microbe. Analyzed here are the complete genomic sequences of *L. pneumophila* Paris (3 503 610 bp; 3 077 genes), an endemic strain predominant in France, and *L. pneumophila* Lens (3 345 687 bp; 2932 genes), an epidemic strain responsible for a major epidemic in France. A striking characteristic of the genome of *L. pneumophila* is its plasticity. Three different plasmids were identified, and ~13 % of each genome is different to the other strain. The Paris strain codes for a unique secretion system of type V, and its secretion system Lvh of type IV is coded by a region of 36 kb which can be either carried on a multicopy plasmid or be integrated into the chromosome. The genetic mobility can be a mechanism which increases the multiplicity of *L. pneumophila*. A large number of genes codes for proteins or patterns of eukaryotic type provided to modulate the functions of the host cell to the advantage of the pathogen, comprising the repeated sequences of tetratrico peptide, ankyrin, F box, serin-threonin kinase proteins, apyrases and a sphingosine-1-phosphate lyase. Therefore, the genome reflects the history and the lifestyle of *L. pneumophila*, a human pathogen of the macrophages which has co-evolved with soft-water amoeba.

L. pneumophila is the causal agent of the legionellosis, an atypical pneumonia, which can be fatal if it is not rapidly treated[1]. This Gram-negative facultative

intracellular pathogen can at the same time adapt to the aquatic environment and to the intracellular medium of the phagocytary cells of the human host[2]. When inhaled in contaminated aerosols, *L. pneumophila* can reach the alveolae of the lungs where they swamped by macrophages. By opposition to the majority of bacteria, which are destroyed, *L. pneumophila* can multiply within the phagosome of the macrophage and in the end kill off the macrophage, resulting in legionellosis[3].

L. pneumophila and other legionelles are inhabitants of natural aquatic biotopes and artificial aqueous systems, such as the cooling towers of air conditioners[3]. Legionelles were detected by culture in soft-water environments at 40 % and by PCR in soft-water environments up to 80 %, where the bacteria are known to survive and replicate by intracellular means in free living protozoons, often within aquatic biological films[4]. Its capacity for exploiting the base cellular mechanisms of a large spectrum of protozoic eukaryotic hosts likewise allows legionelles to infect human cells[5]. In fact, it was shown that the capacity of *L. pneumophila* to multiply by intracellular means in amoeba contributes to the disease, even though little is known of things about the mechanisms governing the host-microbe interactions. Inversely, the biphasic life cycle of *L. pneumophila*, the changing of replicating parasitic cells into transmissible extracellular forms and the complex regulatory network, which governs these changes, are understood in part[6].

The *Legionella* genre comprises 48 species, but more than 90 % of the cases of clinical legionellosis are caused by *L. pneumophila* even more arresting, up to 84 % are caused by the serogroup 1 of *L. pneumophila*[7]. We have determined the complete genomic sequences of two clinical isolates of the serogroup 1, the Paris and Lens strains, to provide knowledge of the genetic characteristics of *L. pneumophila*, and to identify the properties which were selected in niches specific to the pathogenicity and the life cycle of *L. pneumophila*. The Paris strain is the only endemic strain known to date, accounting for 12.7 % of the cases of legionellosis in France and for 33 % of those occurring in the Paris region[8]. It is associated with the nosocomial and community diseases occurring in the form of epidemics or sporadic cases. From November 2003 to January 2004 the Lens strain caused an epidemic of 86 cases resulting in 17 deaths in the north of France, suggesting that it is particularly efficacious for causing the disease in humans. The genomic comparatives of an endemic isolate and of an epidemic isolate supplying the bases for comprehending the specificity of strain, and can give indices for the particular adaptability and stability of the Paris strain.

Results

General characteristics

The Paris strain and the Lens strain of *L. pneumophila* each contain a circular chromosome of ~3 503 610 bp and ~3 345 687 bp, respectively, with an average G+C content of 38 % (Table XXII, Figure 1, Gen–Bank/EMBL access numbers CR628336, CR628337). *L. pneumophila* Paris separates a plasmid of 131 885 bp and the Lens strain contains another plasmid of 59 832 bp (Gen–Bank/EMBL access numbers CR628338, CR628339. In the chromosome of *L. pneumophila* Paris, 3077 genes were identified and 2 932 in that of *L. pneumophila* Lens. No function was able to be predicted for 42.1 % (1354) of the genes of *L. pneumophila* Paris and 44.1 % (1320) of the genes of *L. pneumophila* Lens, a proportion similar to that found in the majority of the other bacterial genomic sequences. A high proportion of the genes provided (21 % for the Paris strain, 20.4 % for the Lens strain) is unique to the *Legionella* genre and can thus code the specific functions of *Legionella*.

Exploitation and modulation of the functions of the host cell

A fascinating question is to know how *Legionella* to decompose the functions of the host to enter, survive, replicate and evade amoeba or alveolar macrophages. Within its genome *L. pneumophila* codes for an abundance of proteins of eukaryotic type. In effect, 30 have the highest similarity with eukaryotic proteins (Table XXIII) and 32 genes code for proteins with eukaryotic domains implied in protein–protein interactions (Table XXIV). We reveal here proteins provided for diverting eukaryotic regulatory paths or for being secreted in eukaryotic cells, making strong candidates of them for directing the invasion of *Legionella*, for traveling in the host cell, or modulating or being subtracted from functions of the host cell.

The repeated sequences of tetratrico peptide (TPR) are repeated patterns of the 34 degenerated amino–acids present in networks in tandem of 3 to 16 patterns, which form hooks for facilitating the protein-protein interactions. The TPR proteins contribute to control of the cellular cycle, to repression of transcription, to response to stress, to inhibition of the protein kinase, to the transport of mitochondrial and peroxisomal protein and to neurogenesis[9]. The Sel-1 repeated sequences represent a sub–family of TPR sequences. In *L. pneumophila* five proteins containing Sel-1 domains were identified. Two of them (EnhC and LidL) were previously implied in interaction with the host cells or in precocious signaling of events which regulate the scheduling

decisions of *L. pneumophila* in the macrophages[10,11]. As a consequence, the three newly identified proteins are likewise in al likelihood implied in the host–pathogen interactions (Table XXIV).

After internalizing, *L. pneumophila* manipulates the endosome–lysosome degradation path of the host for surviving and replicating within a vacuole derived from the endoplasmic reticulum (ER). A protein of *L. pneumophila*, RalF, thought to contribute to recruitment of the ER contains a eukaryotic domain Sec 7. RalF, a substrate of the secretion system of type IV is required by the regulatory protein ARF for associating with the phagosomes of *L. pneumophila*[12]. The two strains of *L. pneumophila* code for three serin/threonin kinase proteins of eukaryotic type (STPK) (Table XXIV). The multisequence comparisons of the domains of the kinase from the Paris strain of *L. pneumophila* and other prokaryotic and eukaryotic STPK, have revealed that Lpp2626 and Lpp1439 of *L. pneumophila* aggregate in the group of eukaryotic STPK, close to the STPK originating from *Entamoeba histolytica* (Figure 2). *Mycobacterium tuberculosis*, which as *L. pneumophila* blocks the phagosome-lysosome fusion, produces eleven STPK of eukaryotic type[13]. In particular, the STPK PknG of *M. tuberculosis* is an inhibitor of the phagosome–lysosome fusion and a promoter of intracellular survival[14]. The STPK domain of Lpp0267/Lpl0262 is related to the PknG and to the STPK YpkA of *Y. pseudotuberculosis* (Figure 2), an enzyme which is translocated in the eukaryotic cells where it corrects the defenses of the host by interfering with the transduction paths of the eukaryotic signal[15]. This suggests that the STPK of *L. pneumophila* can likewise modulate the transduction mechanisms of the eukaryotic signal and can modify the routing paths of the host cell.

Twenty proteins contain ankyrin domains (Table XXIV), sequences repeated in tandem of around 33 amino acids which represent one of the modular interaction protein–protein patterns, the most current being eukaryotic. To date, the only prokaryotic genomes known for coding large families of proteins of the ankyrin domain are *Coxiella burnetii* and *Wolbachia pipientis*, which code 13 and 23 elements, respectively[16,17]. Similar to *L. pneumophila*, *C. burnetii* is an intracellular pathogen which is extremely well-adapted to life inside the eukaryotic phagolysosome, and *W. pipientis* is an "endosymbiont" parasite living in the reproductive cells of a large variety of arthropods. Therefore, the ankyrin domains can be implied in a common microbial mechanism for manipulating the physiology of the host cell. A possible function of the proteins containing the repeated ankyrin sequences of *L. pneumophila* is to modify the

interactions with the host cytoskeleton, given that it is thought that numerous eukaryotic ankyrin proteins are binders between the membranous proteins and the cytoskeleton[18] and are important for targeting proteins towards the plasmic membrane or towards the endoplasmic reticulum. The ankyrin domains are likewise compounds of transcription regulators, suggesting that they can influence the expression of the genes of the host cell as proposed for *Ehrlichia phagocytophila*[19]. In effect, one of the repeated ankyrin proteins of *L. pneumophila* (Lpp3991/Lpl0559) likewise contains a eukaryotic SET domain, known for binding the chromatin host and for influencing the expression of the genes of the host cell[20]. In opposition to the ankyrin proteins of *W. pipientis*, none of the repeated ankyrin proteins of *L. pneumophila* contains a peptide signal[17]. Instead of this, certain of them could be secreted by means of the secretion system of type IV, a way which is independent of the typical targeting signals.

The final stages in the intracellular life cycle of *L. pneumophila* are to kill and escape from its host cell, a mechanism which is still not understood. One class of proteins which can affect the control of the division of the host cell (Lpp2082/Lpl2072, Lpp2486 and Lpp0233/Lpl10234) separates eukaryotic F box domains, sites of protein-protein interactions. Generally, Lpl2072, Lpp2486 and Lpp0233/Lpl10234) separates eukaryotic are associated with other interaction domains[21]. Similarly, two of the identified F box proteins are associated with a repeated ankyrin sequence or a double-spooling pattern, respectively (Table XXIV). The proteins of the F box, assembled in SCF ubiquitin-ligase complexes, determine which substrates are going to be targets for ubiquitination and subsequent proteolysis by the proteasome. Given that the targeted substrates comprise promoters and inhibitors of the cellular cycle as well as transduction compounds of the signal[22], the F box proteins can regulate the division and cellular differentiation. To our knowledge, the only protein of prokaryotic F box described is VirF of *Agrobacterium tumefaciens,* a protein which, it is thought, interacts with the proteins of the host by means of its F box domain to target it for proteolysis[23]. Another pattern implied in the eukaryotic ubiquitination is the U box pattern. The protein Lpp2887, present in the Paris strain but not in the Lens strain, contains such a pattern. It is apparently the first recognized in a prokaryotic organism.

Additional proteins of eukaryotic type identified in the genome of the two strains of *L. pneumophila* are sphingosine-1-phosphate lyase (Lpp2128/Lpl2102) and two secreted apyrases (Lpp1033, Lpp1880/Lpl1000, Lpl1869), suggesting that *L. pneumophila* modulates the cycle of the host cell to its advantage. The broadly

expressed lyase sphingosine-1-phosphate enzyme catalyses the essentially irreversible splitting of the sphingosine-1-phosphate signaling molecule, a product of the degradation of sphingomyeline which regulates cellular proliferation and cellular death in the eukaryote. In effect, the overexpression of the sphingosine-1-phosphate lyase can induce apoptosis in the eukaryote, identifying this enzyme as a double modulator of the metabolism of the sphingosine-1-phosphate and of the ceramid, as well as a regulator of the decisions of cellular direction[24]. In addition, the sphingosine-phosphate plays a central rôle in the development of the *Dictyostelium discoideum* amoeba, given that interruption to the gene results in aberrant distribution of the actin, an abnormal morphogenetic phenotype and a viability occurring during the stationary phase[25]. The two apyrases of *L. pneumophila* are the only ones identified in the prokaryote. The family of apyrase protein comprises enzymes capable of splitting the tri- and diphosphates nucleotides (NTP and NDP) in a calcium- or magnesium-dependent manner. The apyrase was isolated in the autophagy[26] vacuole suggesting that these two proteins influence the destiny of the phagosome of *L. pneumophila* in diminishing the concentration of NTP and NDP during parasitism of the cell.

246 proteins (7.6 %) in the Paris strain and 231 proteins (7.7 %) in the Lens strain have likewise been identified with double-spooling domains provided (CC), many of which likewise show slight similarities with eukaryotic proteins. The CC domains facilitate the protein-protein interactions either for the multimerisation of protein or the macromolecular recognition. Therefore, double-spooling domains can target proteins at the appropriate locale in the eukaryotic host. Interestingly, all the new substrates (SidA-H, SdeC) of the secretion system of type IV identified by Luo and Isberg[27] as well as LidA, LepA and LepB contain double-spooling domains.

To affect the eukaryotic cell, *L. pneumophila* must translocate these proteins of eukaryotic type to the host cytoplasm. As a consequence these proteins are candidate substrates for the secretion system of type IV, as shown for VirF in *A. tumefaciens*[28] or RalF in *L. pneumophila*[12].

Secretion system

At the centre of the pathogenesis of *L. pneumophila* are the loci dot/icm, which together direct the assembling of secretion apparatus of type IV. Even though the two strains contain the complete loci dot/icm, their sequences have variations. In effect, the comparison of the sequence of dot/icm genes of 18 different strains of *L. pneumophila*

has identified a wide range of variations in sequence, and has placed the strains in seven phylogenetic different groups[29]. However, no correlation with the virulence is apparent.

A novel factor of putative virulence of *L. pneumophila*, limited to the Paris strain, is a provided auto–transporter protein. Lpp0779 contains several marks of secretion systems of type V comprising a peptide of N–terminal head for secretion across the internal membrane and a specialized C–terminal domain which forms a pore in the other membrane through which the domain clone passes to the surface of the cell[30]. Its domain clone is a compound of known repeated sequences of haemagglutinin to be implied in the cell–cell aggregation and extremely similar to those of the auto–transporter of AIDA-I and Ag43 of *Escherichia coli*, two proteins implied in the virulence. The bacterial surface protein AIDA-I facilitates adherence to the mammal cells[31], while Ag43 imparts not only a low level of adhesion to certain mammal cells, but likewise facilitates auto–aggregation which is important for the formation of biological film[32]. In a similar way, the auto–transporter of *L. pneumophila* can be implied in adhesion to the host cell and the formation of biological film. In opposition to AIDA–I and Ag43, the auto-transporter of *L. pneumophila* does not have an RGD pattern implied in the bond with the human integrins, and can thus have another interaction domain. The auto–transporter was acquired in the same way by horizontal gene transfer as suggested from its numerous IS upstream and from GC contents of 41 % which exceed the average of the genome of 38 %. Studies of the distribution of this gene in clinical and environmental strains of *L. pneumophila* in combination with the study of its function must provide knowledge of its importance.

In addition, the two strains of *L. pneumophila* contain a secretion method by translocation (Tat) with combined arginine (TatAB and TatC) and completes the secretion systems of type I and II. The system of type II coded by the genes *lspA, lspD-J* is required for the secretion of several enzymes such as lipase A and B (Lpp0533, Lpp1159/Lpl0509, Lpl1164), phosphatase acid Map and SurE (Lpp1120, Lpp1245/Lpl1124, lpl1245), lysophospholipase A (Lpp2291/Lpl2264) and phospholipase PlaB (Lpp1568/Lpl422), proteins which are all present in the two strains.

Metabolism

The metabolic paths utilized by *L. pneumophila* for multiplying inside the eukaryotic cells are not known. The bacteria seems to prefer the proteic substrates, given that a large number of absorption and degradation systems of oligopeptide and

amino acid are coded in the genome. In particular, apart from the elastase homologue ProA of *Pseudomonas aeruginosa* (Lpp0532/lpl0508), three secreted paralogue metalloproteases as well as 46 additional peptidases are present.

By way of opposition, systems for the absorption of sugar are rare, even though the complete ways of Embden-Meyerhof and Entner-Doudoroff are present. In the two strains of *L. pneumophila* no absorption system of type PTS was identified. However, certain of the transport systems of type 55 ABC can be implied in the absorption of sugar, given that the bacteria has some systems for the degradation of complex sugars, such as trehalase, polysaccharide deacetylase, glucoamylase of type eukaryotic (Lpp0489/Lpl0465), β-hexosaminidase and chitinases (Lpp1117/Lpl1121). The two strains code for proteins highly homologous to transporters of glycerol phosphate ABC (Lpp1696, Lpp1695, Lpp1694/Lpl1695, Lpl1694, Lpl1693), and for a hexose phosphate transporter (Lpp2623/Lpl2474), which can be important during intracellular growth. We have likewise identified several enzymes probably implied in the utilization of meso-inositol, which can interfere with the signaling of the host cell facilitated by this intracellular messenger.

L. pneumophila is provided for coding for an extensive aerobic respiratory chain constituted by NADH deshydrogenase, cytochrome-dependent succinate deshydrogenase, ubiquinol-cytochrome reductase and four terminal oxydases, which guarantee the capacity to adapt to changing oxygen tensions (one cytochrome aa_3, two quinol-cytochromes of type *bd* and one quinol cytochrome oxidase of type *o*). The latter oxidase is absent in the Lens strain. Systems implied in anaerobic respiration are apparently absent in all the strains. The two genomes code for an ATP synthase of type F_oF_1 typical of γ-proteobacteria, whereas the Paris strain codes for a second ATP synthase similar to the non-characterized systems of archeobacteria and bacteria marines. *L. pneumophila* likewise codes for at least four sodium/proton anti–carriers (Lpp1464, Lpp2448, Lpp0868, Lpp0667/Lpl1519, Lpl2304, Lpl0839, Lpl0651), which modulate presumably the proton and sodium gradient across the cytoplasmic membrane. As a consequence, a sodium motor force can be utilized for the cellular activities. In this respect, the presence of a compound of polar flagellar motors of type sodium, MotY, thus two significantly different aggregates of gene MotA–MotB, leads to the prediction that mobility can be activated by the sodium motor forces as well as the proton forces. One particular characteristic of *L. pneumophila* is differentiation in a mature intracellular form which accumulates inclusions of poly-hydroxybutyrate in the form of

carbon and energy reserve. As a consequence, the Paris strain codes for four paralogue poly-hydroxybutyrate synthases and the Lens strain codes for three paralogue poly-hydroxybutyrate synthases (Lpp2323, Lpp2038, Lpp2214, Lpp0650/Lpl1055a-b, Lpl2186, Lpl0634).

Physiological adaptation and regulation of the gene

In accord with this intracellular life style the regulatory repertoire is rather small. Analysis of the genome has identified 92 transcription regulators (79 in the Lens strain), which represent only 3.0 % of the genes provided. *L. pneumophila* codes for six sigma putative factors, the homologous of *rpoD*, *rpoH*, *rpoS*, *rpoN*, *fliA* and the sigma factor *rpoE* of type ECF. The number of systems with two compounds (13 histidine kinases and 14 response regulators) is likewise low.

The most abundant class of regulators belongs to the GGDEF/EAL (23) family. Present in numerous bacteria, comprising *Vibrio cholerae* (41), *P. aeruginosa* (33), *Wolinella succinogenes* (26), and *E. coli* (19), these regulators contain two sub-domains, GGDEF and EAL. Of the 23 regulators identified, 10 separate only one GGDEF domain, 3 in the Paris strain and 2 in the Lens strain contain an EAL domain, and 10 in the Paris strain and 11 in the Lens strain present a combination of the two. The rôle of these regulators in *L. pneumophila* is unknown, but in other bacteria these regulators play a rôle in aggregation, the formation of biological film or mobility by muscular contraction.

In *L. pneumophila*, the cyclic AMP can likewise translate the cellular signals given that the genome codes for five adenylate cyclases of class III (Lpp1446, Lpp1131, Lpp1704, Lpp1277, Lpp0730/Lpl1538, Lpl1135, Lpl1703, Lpl1276, Lpl0710). In *P. aeruginosa*, CyaB, an adenylate cyclase of class III, is implied in the regulation of the virulence of gene 33. However, *L. pneumophila* does not contain the orthologue of Vfr, the dependent AMPc regulator of *P. aeruginosa*, but it codes for five proteins with AMPc bonding patterns (Lpp3069, Lpp1482, Lpp2063, Lpp0611, Lpp2777/Lpl2926, Lpl1501, Lpl2053, Lpl0592a-b-c, Lpl2648). As for *P. aeruginosa*, these adenylate cyclases of class III can comprise environmental signals extending from the nutritional content of the surrounding medium to the presence of host cells and can control the expression of the virulence of the gene in consequence.

Heightened plasticity of the genomes of *L. pneumophila*

The two genomes exhibit an astonishingly high plasticity and a diversity of the genome. Comparison of the chromosomes has identified a preserved skeleton of 2664 genes but 280 and 428 genes (10 and 14 %) are specific to the strain (Figure 3). Given that the two strains analyzed belong to the same species and the same serogroup, this diversity was unexpected. For example the comparison of the genomes of two strains of *Salmonella enterica* of serotype Typhi identifies only 2 % for each of the genes specific to the strain[34]. The specific genetic equipment of the Paris strain contains a certain number of regulators (three homologous of CsrA, 13 transcription regulators), of additional repeated ankyrin sequences and proteins of type eukaryotic (Table XXIII and Table XXIV) as well as several restriction modification genes (modification methylases of the DNA, endonucleases), which can explain the low competence (personal observation) and the high genomic stability[8] of the Paris strain. The Lens strain contains fewer specific regulators (4) and four specific proteins with eukaryotic domains (Table XXIII), two of which are repeated sequences of ankyrin proteins, suggesting that the Paris strain is a particularly well-equipped strain.

The genomes of *L. pneumophila* have undergone rearrangements of multiple genomes. The important synteny in the genome between the Paris and Lens strain is interrupted by inversion of 260 kb, insertion of 130 kb in the Paris strain (or deletion in the Lens strain) and by deletions and smaller multiple insertions. The fragment of 130 kb is flanked by an ARnt gene and codes for a putative integrase, suggesting a structure similar to the islets of pathogenicity of the enterobacteria. It contains an ATP synthase, chemiosmotic flow systems (cebABC, cecABC) and the genes cadA1, ctpA, copA1, copA2 coding for ATP-dependent flow pumps, as was proven to be induced in the macrophages[35] and separate the prpA-lvrABC gene aggregate, present within a pathogenicity islet of 65 kb in the Philadelphia strain[36]. With the exception of abovementioned genes, this pathogenicity islet is absent from the Paris strain and from the Lens strain. However, the corresponding chromosomal site in the Paris strain is the insertion site of an integrative plasmid discussed below. Therefore, these two regions can be hot points for genomic rearrangements. Genomic variation is likewise evident from its network of mobile elements comprising ten integrases, 58 insertion sequences (34 complete and 24 truncated) thus as proteins relevant to phages. In addition, the genomes contain a large number of repeated sequences organized in the form of repeated inverse sequences, which recall the ERIC sequences of the enterobacteria.

These LeRIC (Repeated Intergenic Consensus of Legionella) fall into 7 classes present in numerous copies (for example 80, 18, 18, 25, 9, 9 and 6 in the Paris strain).

L. pneumophila Paris and Lens contain lvh, a region which codes for a second secretion system of type IV previously characterized in the Philadelphia strain[37]. One interesting observation is that the lvh region of *L. pneumophila* Paris is coded on a region of 36 kb which exists either integrated into the chromosome or excised in the form of a multicopy plasmid (unpublished data). This pattern is similar to that described for the unstable element of 30 kb of the Olda strain, which is possibly phage-derived and is implied in the phase variation[38]. The GC content of the lvh region (43 %) is different to that of the remainder of the chromosome (38 %) and it contains certain genes related to phages, suggesting possible phagic origin. However, the exact excision and integration mode is still not understood. An attractive hypothesis is that the integration and excision of particular regions of the chromosome is a mechanism specific to *L. pneumophila* for boosting versatility.

The second plasmid of the Paris strain (132 kb) comprises known virulence factors, mobile genetic elements and genes of antibiotic resistance. The regulator system with two lrpR–lskS compounds present on this plasmid was found on a plasmid of *Legionella longbeachae* implied in the virulence of these species[39]. Heightened preservation (93 to 98 % of protein identity) of the six gene sequences on the plasmid of 135 kb along *L. longbeachae* can indicate a recent horizontal transfer between *L. pneumophila* and *L. longbeachae*. *L. pneumophila* Lens contains a plasmid of 60 kb which codes for several proteins homologous to the transfer region of the F plasmid of *E. coli*. All three plasmids of the Paris and Lens strain code for a paralogue of CsrA, a repressor of the transmission traits and activator of replication[40].

Although the rôle of the plasmids in the virulence of *L. pneumophila* has yet to be determined, the correlation between strains containing a plasmid and the virulence in a mouse model was described[41]. In addition, *L. pneumophila* strains with plasmids seem to persist longer in the environment than those strains not having plasmids[42]. The identification in the clinical isolates of the plasmids coding for factors of putative virulence is another indication of the importance of the plasmids for the pathogenicity of *Legionella*.

The genomes of *L. pneumophila* display a plasticity likewise at the gene level. The loci *enh*, implied in the entry in the host cells[43], are present in the Paris and Lens strains. One of the proteins coded by these loci is RtxA, which contributes to entry,

adherence, cytotoxicity, pore formation[43] and intracellular routing in the amoeba[44]. Unlike the AA100[10] strain in the Paris strain, rtxA is fused with arpB and a second protein with about 30 sequences repeated in tandem extremely preserved de 549 bp. A similar structure is coded by the Lens strain, however we have identified two patterns within the repeated region, the two being different to that of the Paris strain. However, the number of repetitions seems to be the same (Figure 4). It is possible that the variations in number and sequence of the repeated sequences contribute to the multiplicity and likewise to the virulence of *L. pneumophila*.

In accordance with the relative plasticity in their genome, the strains of *L. pneumophila* code for organelles of type IV bacterial pili, which are required for natural competence for the transformation of the DNA[45]. The organization of the genes coding for type IV bacterial pili is similar to that of *P. aeruginosa* where they are crucial for bacterial adherence and colonization of the mucosal surfaces and for mobility by muscular contraction. Another mechanism in *L. pneumophila* contributing to the plasticity of the genome is conjugated transfer facilitated by the secretion of type IV of plasmids[46] and chromosomal DNA[47].

Conclusion

Analysis of the sequences of the genome of the clinical of *L. pneumophila* Paris and Lens strains and its comparison identifies *L. pneumophila* as an extremely versatile organism, which demonstrates a plasticity and an extensive genomic diversity. The excision and integration of plasmids or genes can be a mechanism which *L. pneumophila* exploits for adapting to different environments. Its large cohort of proteins of eukaryotic type is provided for manipulating the host cell to the advantage of the pathogen (Figure 5). Eucaryotic proteins of putative origin have likewise been identified in other intracellular pathogens, comprising *Coxiella*, *Wolinella*, *Agrobacterium*, *Mycobacterium* and *Ehrlichia*, but currently *L. pneumophila* sequences them as prokaryotic with the greatest variety of proteins of eukaryotic type. Presumably, during its co-evolution with free living amoeba, the *L. pneumophila* pathogen acquires DNA by horizontal transfer from its host or by convergent evolution. These proteins can then likewise contribute to the infection of human macrophages. By being based on the genomic sequences future comparative and functional studies are going to enable survival tactics of intracellular parasites to be defined, and to identify the special attributes of endemic and epidemic *L. pneumophila*. To combat the menace of *L.*

pneumophila resistant to chemical products widely utilized for decontamining public water systems, comprising hospitals, the genomic sequence can stimulate the identification of targets for novel active biocides against *L. pneumophila*.

Methods

Preparation and sequencing of the DNA. Paris and Lens strains of *L. pneumophila* were cultivated on agar BCYE at 37°C over 3 days and the chromosomal DNA was isolated by utilizing standard protocols. The cloning, sequencing and assembling were completed as described previously[48]. For the two genomes, two libraries (inserts of 1-2 kb and 2-3 kb) were generated by random mechanical chiseling of the genomic DNA and cloning in pcDNA-2,1 (Invitrogen). A hook was obtained by terminal sequencing of clones from a BAC library constructed as described previously[49] by utilizing pIndigoBac (Epicentre) as vector. For the Paris strain of *L. pneumophila* an insert library of average size (8-10 kb) was constructed in the low-number vector of copies pSYX34. The purification Plasmide DNA was produced either with Montage Plasmide Miniprep96 Kit (Millipore) or by utilizing the DNA sequencing matrix amplification kit TempliPhi (Amersham Biosciences). The sequencing reactions were created by utilizing the sequencing reactions kit ABI PRISM BigDye Terminator and an analyzer 3700 or a 3730 Xl Genetique Analyzer (Applied Biosystems). 47,200 sequences for the Paris strain of *L. pneumophila*, and 47,231 sequences for the Lens strain each originating from four libraries were obtained and assembled and finished as described previously 48.

Annotation and analysis. The definition and annotation of the coding sequences (CDS) is as was described previously[48] by utilizing the Boîte CAAT 50 software. All the CDS provided were checked visually. The predictions on function were based on preferred BLASTp similarity and on analysis of patterns by utilizing the PFAM, Prosite and SMART databases. We have identified orthologous genes by better BDERNIÈRE reciprocal correlation and FASTA comparisons. For identification of the double-spooling domains the publicly available software PairCoil and Coilscan were utilized. The pseudogenes have one or more mutations which prevent complete translation. Repetitive sequences of DNA were identified by BLASTN comparisons of the intergenic regions and of the complete genome. MFOLD software was utilized to predict the folding of the single sheet of DNA molecules.

URL. The sequence and annotation of the two genomes of *L. pneumophila* are at http://genolist.pasteur.fr/LegioList. For annotation and analysis we used PairCoil http://paircoil.lcs.mit.edu/cgi–bin/paircoil and Coilscan http://www.biology.wustl.edu/gcg/coilscan.html

Bibliography for Example 7:

1. McDade, J.E. *et al.* Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. N. Engl. J. Med. 297, 1197-1203 (1977).

2. Fields, B.S. The social life of Legionellae. in Legionella (ed. Marre R., Kwaik, Y.A., Bartlett, C., Cianciotto, N.P., Fields, B.S., Frosch, M., Hacker, J. & Lück, P.C.) 135-142 (ASM Press, Washington D.C., 2002).

3. Steinert, M., Hentschel, U. & Hacker, J. Legionella pneumophila: an aquatic microbe goes astray. FEMS Microbiol. Rev. 26, 149-162 (2002).

4. Fields, B.S., Benson, R.F. & Besser, R.E. Legionella and Legionnaires' disease: 25 years of investigation. Clin. Microbiol. Rev. 15, 506-526 (2002).

5. Atlas, R.M. Legionella: from environmental habitats to disease pathology, detection and control. Environ. Microbiol. 1, 283-293 (1999).

6. Molofsky, A.B. & Swanson, M.S. Differentiate to thrive: lessons from the Legionella pneumophila life cycle. Mol. Microbiol. 53, 29-40. (2004).

7. Yu, V.L. et al. Distribution of Legionella species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: an international collaborative survey. J. Infect. Dis. 186, 127-128 (2002).

8. Aurell, H. et al. Legionella pneumophila Serogroup 1 Strain Paris: Endemic Distribution Throughout France. J. Clin. Microbiol. 41, 3320-3322 (2003).

9. Goebl, M. & Yanagida, M. The TPR snap helix: a novel protein repeat pattern from mitosis to transcription. Trends Biochem. Sci. 16, 173-177 (1991).

10. Cirillo, S.L.G., Lum, J. & Cirillo, J.D. Identification of a novel loci involved in entry by Legionella pneumophila. Microbiology 146, 1345-1359 (2000).

11. Conover, G.M., Derre, I., Vogel, J.P. & Isberg, R.R. The Legionella pneumophila LidA protein: a translocated substrate of the Dot/Icm system associated with maintenance of bacterial integrity. Mol. Microbiol. 48, 305-321 (2003).

12. Nagai, H., Kagan, J.C., Zhu, X., Kahn, R.A. & Roy, C.R. A bacterial guanine nucleotide exchange factor activates ARF on Legionella phagosomes. Science 295, 679-682 (2002).

13. Cole, S.T. et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genomic sequence. Nature 393, 537-544 (1998).

14. Walburger, A. *et al.* Protein Kinase G from Pathogenic Mycobacteria Promotes Survival Within Macrophages. Science 304,1800-1804 (2004).

15. Barz, C., Abahji, T.N., Trulzsch, K. & Heesemann, J. The Yersinia Ser/Thr protein kinase YpkA/YopO directly interacts with the small GTPases RhoA and Rac-1. FEBS Lett. 482, 139-143 (2000).

16. Seshadri, R. *et al.* Complete genomic sequence of the Q-fever pathogen Coxiella burnetii. Proc. Natl. Acad. Sci. U S A 100, 5455-5460 (2003).

17. Wu, M. et al. Phylogenomics of the Reproductive Parasite Wolbachia pipientis wMel: A Streamlined Genome Overrun by Mobile Genetic Elements. PLoS Biol 2, E69 (2004).

18. Hryniewicz-Jankowska, A., Czogalla, A., Bok, E. & Sikorsk, A.F. Ankyrins, multifunctional proteins involved in many cellular pathways. Folia Histochem. Cytobiol. 40, 239-249 (2002).

19. Caturegli P, A.K., Walls, J.J., Bakken, J.S., Madigan, J.E., Popov, V.L., Dumler, J.S. ankA: an Ehrlichia phagocytophila group gene encoding a cytoplasmic protein antigen with ankyrin repeats. Infect. Immun. 68, 5277-5283 (2000).

20. Jenuwein, T., Laible, G., Dorn, R. & Reuter, G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell. Mol. Life Sci. 54, 80-93 (1998).

21. Kipreos, E.T. & Pagano, M. The F-box protein family. Genome Biol. 1, Epub 2000 Nov 10 (2000).

22. Craig, K.L. & Tyers, M. The F-box: a new pattern for ubiquitin dependent proteolysis in cell cycle regulation and signal transduction. Prog. Biophys. Mol. Biol. 72, 299-328 (1999).

23. Schrammeijer, B. *et al.* Interaction of the virulence protein VirF of Agrobacterium tumefaciens with plant homologs of the yeast Skp1 protein. Curr. Biol. 11, 258-262 (2001).

24. Reiss, U. et al. Sphingosine-phosphate lyase enhances stress-induced ceramide generation and apoptosis. J. Biol. Chem. 279, 1281-1290 (2004).

25. Li, G., Foote, C., Alexander, S. & Alexander, H. Sphingosine-1-phosphate lyase has a central role in the development of Dictyostelium discoideum. Development 128, 3473-3483 (2001).

26. Biederbick, A., Rose, S. & Elsasser, H.P. A human intracellular apyrase-like

protein, LALP70, localizes to lysosomal/autophagic vacuoles. J. Cell. Sci. 112, 2473-2484 (1999).

27. Luo, Z.Q. & Isberg, R.R. Multiple substrates of the Legionella pneumophila Dot/Icm system identified by interbacterial protein transfer. Proc. Natl. Acad. Sci. U S A 101, 841-846 (2004).

28. Vergunst, A.C. *et al.* VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science 290, 979-982 (2000).

29. Morozova, I. *et al.* Comparative sequence analysis of the icm/dot genes in Legionella. Plasmid 51, 127-147 (2004).

30. Desvaux, M., Parham, N.J. & Henderson, I.R. The autotransporter secretion system. Res. Microbiol. 155, 53-60 (2004).

31. Benz, I. & Schmidt, M.A. AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic Escherichia coli strain 2787 (O126:H27), is synthesized via a precursor molecule. Mol. Microbiol. 6, 1539-1546 (1992).

32. Klemm, P., Hjerrild, L., Gjermansen, M. & Schembri, M.A. Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from Escherichia coli. Mol Microbiol. 2 51, 283-296 (2004).

33. Smith, R.S., Wolfgang, M.C. & Lory, S. An adenylate cyclase-controlled signaling network regulates Pseudomonas aeruginosa virulence in a mouse model of acute pneumonia. Infect. Immun. 72, 1677-1684 (2004).

34. Deng, W. et al. Comparative Genomics of Salmonella enterica Serovar Typhi Strain. J. Bacteriol. 185, 2330-2337 (2003).

35. Rankin, S., Li, Z. & Isberg, R.R. Macrophage-induced genes of Legionella pneumophila: protection from reactive intermediates and solute imbalance during intracellular growth. Infect. Immun. 70, 3637-3648 (2002).

36. Brassinga, A.K. *et al.* A 65-Kilobase Pathogenicity Island Is Unique to Philadelphia-1 Strains of Legionella pneumophila. J. Bacteriol. 185, 4630-3637 (2003).

37. Segal, G., Russo, J.J. & Shuman, H.A. Relationships between a new type IV secretion system and the icm/dot virulence system of Legionella pneumophila. Mol. Microbiol. 34, 799-809 (1999).

38. Luneberg, E. *et al.* Chromosomal insertion and excision of a 30 kb unstable genetic element is responsible for phase variation of lipopolysaccharide and other virulence determinants in Legionella pneumophila. Mol. Microbiol. 39, 1259-1271 (2001).

39. Doyle, R.M. & Heuzenroeder, M.W. A mutation in an ompR-like gene on a Legionella longbeachae serogroup 1 plasmid attenuates virulence. Int. J. Med. Microbiol. 292, 227-239 (2002).

40. Molofsky, A.B. & Swanson, M.S. Legionella pneumophila CsrA is a pivotal repressor of transmission traits and activator of replication. Mol. Microbiol. 50, 445-461 (2003).

41. Bezanson, G., Fernandez, R., Haldane, D., Burbridge, S. & Marrie, T. Virulence of patient and water isolates of Legionella pneumophila in guinea pigs and mouse L929 cells varies with bacterial genotype. Can. J. Microbiol. 40, 426-431 (1994).

42. Brown, A., Vickers, R.M., Elder, E.M., Lema, M. & Garrity, G.M. Plasmid and surface antigen markers of endemic and epidemic Legionella pneumophila strains. J. Clin. Microbiol. 16, 230-235 (1982).

43. Cirillo, S.L., Bermudez, L.E., El-Etr, S.H., Duhamel, G.E. & Cirillo, J.D. Legionella pneumophila entry rtxA gene is involved in virulence. Infect. Immun. 69, 508-517 (2001).

44. Cirillo, S.L., Yan, L., Littman, M., Samrakandi, M.M. & Cirillo, J.D. Role of the Legionella pneumophila rtxA gene in amoebae. Microbiology 148, 1667-1677 (2002).

45. Stone, B.J. & Kwaik, Y.A. Natural competence for DNA transformation by Legionella pneumophila and its association with expression of type IV pili. J. Bacteriol. 181, 1395-1402 (1999).

46. Vogel, J.P., Andrews, H.L., Wong, S.K. & Isberg, R.R. Conjugative transfer by the virulence system of Legionella pneumophila. Science 279, 873-876 (1998).

47. Miyamoto, H., Yoshida, S., Taniguchi, H. & Shuman, H.A. Virulence conversion of Legionella pneumophila by conjugal transfer of chromosomal DNA. J. Bacteriol. 185, 6712-6718 (2003).

48. Glaser, P. *et al.* Comparative genomics of Listeria species. Science 294, 849-852 (2001).

49. Buchrieser, C. *et al.* The 102-kilobase pgm locus of Yersinia pestis: sequence analysis and comparison of selected regions among different Yersinia pestis and Yersinia pseudotuberculosis strains. Infect. Immun. 67, 4851-4861 (1999).

50. Frangeul, L. *et al.* CAAT-Box, Contigs-Assembly and Annotation tool-box for genome sequencing projects. Bioinformatics 20, 790-797 (2004).

Table XXII : General characteristics of the two *Legionella pneumophila* genomes

	L. pneumophila Paris	*L. pneumophila* Lens
Size of the chromosome (bp)	3 503 610	3 345 687
G+Ct content	38.3% (37.4%)	38.4% (38.4%)
G+C protein-coding genes content	39.1% (37.9%)	39.4 %(39.1%)
Total number of protein-coding genes	3076 (141)	2931 (57)
Average length (codons) of protein-coding genes	331	333
Number of operons rRNA (16S-23S-5S)	3	3
Number of genes tRNA	43	43
Coding in percentage	87.9% (92%)	88% (83.7%)
Plasmid	1 (131.9 kb)	1 (59.8 kb)
Number of genes specific to the line	428 (125)	280 (44)
Number of orthologous genes	2664	2664

[1]kilobase pairs

Table XXIII: Proteins having the greatest similarity with eukaryotic proteins

***L. pneumophila* Paris**	**Produit provided**	***L.p.*Lens**	**G-C**	**Percentage of protein identity**
lpp1647	purC	*lpl1640*	38%	61% over the entire length (AAR06292.1 *Nicotania tabacum*)
lpp0702	exoA exodeoxyribonuclease III	*lpl0684*	39%	58% over the entire length (EAA20230.1 *Plasmodium yoelii yoelii*)
lpp0321	Precursor protein bond to RNA	-	34%	45% of 50% of the protein (AAL07519 *Solanum tubeosum*)
lpp1157	pyruvate decarboxylase	*lpl1162*	39%	50% over the entire length (AAB16855.1 *Arabidopsis thaliana)*
lpp1522	protein de biosynthesis de thiamine NMT-1	*lpl1461*	38%	49% over the entire length (AAC64375.1 *Botryotinia fuckeliana*)
lpp2832	nuoE NADH dehydrogenase I chain E	*lpl2701*	38%	49% of 82% of the protein (BAA25988.1 *Homo sapiens*)
plpp0050	protein with hypersensitive induced response	-	36%	48% over the entire length (AAN17462.1 *Hordeum vulgare subsp. Vulgare*)
lpp0634	hypothetical protein	*lpl0618*	39%	48% over the entire length (XP_306643.1 *Anopheles gambiae*)
lpp0965	protease	*lpl0935*	39%	45% over the entire length (NP_189431.2 *Arabidopsis thaliana*)
lpp2748	phytanoyl-coA dioxygenase	*lpl2621*	36%	44% over the entire length (XP_372144.1 *Homo sapiens*)
lpp2128	sphingosine-1-phosphate lyase	*lpl2102*	41%	36% over the entire length (NP_775139.1 *Rattus norvegicus*)

lpp0489	glucoamylase	*lpl0465*	39%	32% over the entire length (P42042 *Arxula adeninivorans*)
lpp0955	cytokinin oxydase	*lpl0925*	39%	32% over the entire length (NP_484368.1 *Nostoc sp.*)
lpp0578	phytanoyl coA dioxygenase	*lpl0554*	36%	31% over the entire length (EAA70100.1 *Gibberella zeae*)
lpp0379	hypothetical protein	*Lpl0354*	39%	31% over the entire length (CAD21525.1 *Taenia solium*)
lpp1033	ectonucleoside triphosphate diphosphohydrolase (apyrase)	*lpl1000*	40%	25% over the entire length - nucleoside phosphatase signature (Q9MYU4 *Sus scrofa*)
lpp2923	6-pyruvoyl-tetrahydropterin synthase	*lpl2777*	34%	26% over the entire length (NP_703938.1 *Plasmodium falciparum*)
lpp3071	zinc metalloproteinase	*lpl2927*	38%	38% over the entire length (AAF56122.1 *Drosophila melanogaster*)
lpp2134	methyltransferase bonded to SAM	*lpl2109*	35%	24% over the entire length (BAC98835.1 *Bombyx mori*)
lpp1880	ectonucleoside triphosphate diphosphohydrolase (apyrase)	*lpl1869*	39%	26% over the entire length (CAE70887.1 *Caenorhabditis briggsae*)
lpp2747	methyltransferase bonded to SAM	*lpl2620*	35%	33% of 56% of the protein (EAA20288.1 *Plasmodium yoelii yoelii*)
lpp2468	Cytochrome P450	*lpl2326*	39%	20% of 75% of the protein (NP_487786.1 *Nostoc sp.*)
lpp1824	protein bound to the nuclear membrane	-	34%	19% of 40% of the protein (NP_082559.1 *Mus musculus*)
lpp1665	uracyl DNA glycosylase	*lpl1659*	36%	27% of 80% of the protein (EAA36774.1 *Giardia lamblia*)
lpp1959	condensation chromosome type 1	*lpl1953*	41%	Model of preserved chromosome condensation regulator
lpp0358	hypothetical protein	*lpl334*	38%	37% on 53% of the protein (EAA20288.1 *Plasmodium yoelii yoelii*)
lpp1127	Ca2+-ATPase de transport	*lpl1131*	37%	22% on 34% of the protein (AAB81284.1 *Paramecium tetraurelia*)
lpp1167	uridine kinase	*lpl1173*	33%	35% on 65% of the protein (AAM09314.2 *Dictyostelium discoideum*)
lpp2626	domain ser/thr kinase protein	*lpl2481*	32%	domain preserved
lpp1439	serine threonine kinase protein	*lpl1545*	36%	domain preserved

Lpp indicates the coding sequences (CDS) provided from *L. pneumophila*strain Paris; *lpl* indicates the coding sequences (CDS) provided from *L. pneumophila*strain Lens, the lines in gray indicate the proteins which are likewise mentioned in Table 2B in terms of their preserved eukaryotic domains; the access numbers to the proteins are indicated between parentheses

Table XXIV: coding domains of *L. pneumophila* preferably protein found in eukaryotic proteins

L. pneumophila Paris	*L. pneumophila* Lens	identified unit	Content G-C	putative function
enhC (lpp2692)	*enhC (lpl2564)*	21 domains sel-1	39%	Invasion and traffic in host cells
lidL(lpp1174)- EnhC paralog	*lidL (lpl1180)* - EnhC paralog	6 domains sel-1	38%	
lpp1310- EnhC paralog	*lpl1307* - EnhC paralog	4 domains sel-1	41%	
lpp2174- EnhC paralog	*lpl1303* - EnhC paralog	3 domains sel-1	40%	
-	*lpl1059* - EnhC paralog	7 domains sel-1	45%	
ralF (lpp1932)	*ralF lpl1919*	domain sec7	34%	
lpp0267	*lpl0262*	domain ser/thr kinase protein	38%	Modulation functions of host cells
lpp2626	*lpl2481*	domain ser/thr kinase protein	32%	
lpp1439	*lpl1545*	domain ser/thr kinase protein	36%	
lpp2065	*lp2055*	ankyrin repetition	37%	
lpp0037	*lpl0038*	ankyrin repetition	38%	
plpp0098	-	ankyrin repetition	37%	
lpp2058	*lpl2048*	ankyrin repetition	38%	
lpp0750	*lpl0732*	ankyrin repetition	35%	
lpp2061	*lpl2051*	ankyrin repetition	39%	
lpp2270	*lpl2242*	ankyrin repetition	34%	
lpp0503	*lpl0479*	ankyrin repetition	38%	
lpp1905	-	ankyrin repetition	35%	
lpp1683	*lpl1682*	ankyrin repetition + domain SET	33%	
lpp2248	*lpl2219*	ankyrin repetition	39%	
lpp0202	-	ankyrin repetition	38%	
lpp0469	*lpl0445*	ankyrin repetition	38%	
lpp2517	*lpl2370*	ankyrin repetition	36%	
lpp1100	-	ankyrin repetition	48%	
lpp0126	*lpl0111*	ankyrin repetition	39%	
lpp0356	-	ankyrin repetition	38%	
lpp2522	*lpl2375*	ankyrin repetition	39%	
lpp0547	*lpl0523*	ankyrin repetition	40%	
-	*lpl1681*	ankyrin repetition	34%	
-	*lpl2344*	ankyrin repetition	35%	
-	*lpl2058*	ankyrin repetition	40%	
lpp2082	*lpl2072*	domain F-box + ankyrin repetition	36%	Control division, evasion host cells
lpp2486	-	domain F-box + superhelice	34%	
lpp0233	*lpl0234*	domain F-box	39%	
lpp2887	-	2 domains U-box	35%	
lpp2128 Sphingosine-1-phosphate lyase	*lpp2102* Sphingosine-1-phosphate lyase		41%	

Lpp indicates the coding sequences (CDS) provided from *L. pneumophila*strain Paris; *lpl* indicates the coding sequences (CDS) provided from *L. pneumophila*strain Lens; the numbers indicate the number of domains identified in a protein; ser/thr indicates threonine serin

Example 9: Repeated sequences of DNA and secreted enzymes characteristic of Legionella pneumophyla

A) Repeated sequence

A repeated sequence was identified in the genome of the Legionella pneumophila Paris strain then sequenced completely. This sequence (SEQ ID 7074) is of 122 bp and is repeated 86 times in the genome of the L. pneumophila Paris strain. The preservation is from 81 to 100 % (0 to 19 mismatch) over the entire length for 53 copies and from 70 to 80 % over a length of at least 100 nucleotides for 33 copies. In the L. pneumophila Lens strain, there are 62 copies whereof the preservation of 29 copies is from 81 % to 95 % over the entire length. We have determined oligonucleotides specific to this sequence for its amplification by PCR. Tests on 15 strains each belonging to one of the serotypes of L. pneumophila (serotypes 1 to 15) have shown that this sequence is present in all serotypes of L. pneumophila. The test of 11 strains belonging to other species of the Legionella genre (L. miedadei, L. dumoffii, L. gormanii, L. longbeachae serogroup 1, L. jordanis, L. anisa, L. erythre, L. rubriluccns, L. quinlivani, L. moravica, L. taurinensis) with these oligonucleotides, thus databank research have shown that this sequence is specific to the Legionella pneumophila species and that it will be able to thus serve as an identifier of the species (see Tables XXV and XXVI and Figures 6 and 7). The high number of copies of this sequence in the genome will enable amelioration of the sensitivity of a PCR test or by hybridization, compared to a present sample in a single copy.

SEQ ID 7074:

AGGACTTACGAAAAACCCCAAGATCAAGGCAAAAAATGTTTTTAATGAGG
GAGTTTAGATAAACTAAATAACCGAATTAAAAACTTTTTTTAACAAAGAG
ATTGGGATTTTTCGTAAGTCCT.

This sequence is an interesting target for diagnostics of L. pneumophila diagnostic by PCR or by other methods equivalent to PCR.

Among the primers used for amplification and detection of this repeated sequence, the following couple of primers can especially be cited:

SEQ ID N° 7075: GAAAAACCCCAAGATCAAGGC and

SEQ ID N° 7076: AGGACTTACGAAAAACCCCAA.

Table XXV: List of the DNA of the 15 reference serogroups of *Legionella pneumophila*

Name	Number labo	n°ATCC
L pneumophila **sg1**	R1	ATCC 33152
L pneumophila **sg3**	R2	ATCC 33155
L pneumophila **sg2 Togus 1**	R6	ATCC 33154
L pneumophila **sg4** LosAngeles1	R7	ATCC 33156
L pneumophila **sg5** subsp fraseri	R8	ATCC 33216
L pneumophila **sg6** Chicago 2	R10	ATCC 33215
L pneumophila **sg7** Chicago 8	R12	ATCC 33283
L pneumophila **sg8** Concord 3	R13	ATCC 33096
L pneumophila **sg9** IN-23-G1-E2	R14	ATCC 35289
L pneumophila **sg10** Leiden 1	R15	ATCC 43283
L pneumophila **sg11** 797-PA-H	R18	ATCC 43130
L pneumophila **sg12** 570-CO-H	R19	ATCC 43290
L pneumophila **sg13** 82A3105	R20	ATCC 43736
L pneumophila **sg14** 1169-MN-H	R21	ATCC 43073
L pneumophila **sg15** Lansing3	R22	ATCC 35251

Table XXVI : list of the DNA of the 11 reference species *Legionella* non *pneumophila*

ATCC Number	Description	Strain Reference
33218	*Tatlockia micdadei* Garrity *et al.* deposited as Legionella micdadei	TATLOCK [CIP 103882; NCTC 11371]
33279	*Fluoribacter dumoffii* (Brenner et al.) Brown *et al.* deposited as Legionella dumoffii	NY 23
33297	*Fluoribacter gormanii* (Morris *et al.*) Brown *et al.* deposited as Legionella gormanii	LS-13 [ALLO3]
33462	*Legionella longbeachae* McKinney *et al.* serogroup 1	Long Beach 4 [NCTC 11477]
33623	*Legionella jordanis* Cherry *et al.*	BL-540
35291	*Legionella anisa Gorman et al.*	CH-47-C3
35303	*Legionella erythra* Brenner *et al.*	SE-32A-C8 [NCTC 11977]
35304	*Legionella rubrilucens Brenner et al.*	WA-270A-C2 [NCTC 11987]
43830	*Legionella quinlivanii Benson et al.* serogroup 1	1442-AUS-E [CIP 105272]
43877	*Legionella moravica Wilkinson et al*	316-86 [CDC 1634-CZK-E; CIP 103883]
700508	*Legionella taurinensis* Lo Presti *et al.*	Turin I no 1

B) Enzymes secreted

The enzymes secreted and common to the three strains of L. pneumophila (Paris, Lens and Philadelphia) whereof the sequences are identified hereinbelow can be utilized especially as a target in colorimetric tests (or for their being made available) for detection of the presence or not of Legionella in a biological sample.

- the sequence lpp0489 (SEQ ID 4292) which codes for a precursor of glucoamylase (Glucan 1,4-alphä-glucosidase) of eukaryotic cell without homologue in the bacteria;
- the sequence lpp1117 (SEQ ID 6477) which codes for a potential secreted chitinase; and

- the sequences lpp1033 (SEQ ID 4267) and lpp1880 (SEQ ID 3675) which code for a protein similar to an ectonucleoside triphosphate diphosphohydrolase (apyrase) secreted from a eukaryotic cell.

Bibliography (beyond Example 8)

1. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. [Review] [90 refs]. *Nucleic Acids Research.* 25:3389-402.

2. Aurell, H., J. Etienne, F. Forey, M. Reyrolle, P. Girardo, P. Farge, B. Decludt, C. Campese, F. Vandenesch and S. Jarraud. 2003. *Legionella pneumophila* serogroup 1 strain Paris: Endemic distribution throughout France. J Clin Microbiol. 41: 3320-3322

3. Birnboim, H. C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.* 100:243-255.

4. Buchrieser, C., C. Rusniok, L. Frangeul, E. Couve, A. Billault, F. Kunst, E. Carniel, and P. Glaser. 1999. The 102 kb locus of *Yersinia pestis* : sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect. Immun.* **67**:4851-4861.

5. Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8:186-194.

6. Fitch, W. S. 1970. Distingishing homologous from analogous proteins. *Syst. Zool.* 19:99-113.

7. Frangeul L, Glaser P, Rusniok C, Buchrieser C, Duchaud E, Dehoux P and Kunst F. 2004. CAAT-Box, Contigs-Assembly and Annotation tool-box for genome sequencing projects. *Bioinformatics.* in press.

8. Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8:195-202.

10. Li, P., K. C. Kupfer, C. J. Davies, D. Burbee, G. A. Evans, and H. R. Garner. 1997. PRIMO: A primer design program that applies base quality statistics for automated large-scale DNA sequencing. *Genomics.* 40:476-485.

11. Lukashin, A. V., and M. Borodovsky. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 15:1107-1115.

12. Xu, S.Y., and Fomenkov, A. 1994. Construction of pSC101 derivatives with Camr and Tetr for selection or LacZ for blue/white screening. *Biotechniques* **17**:57.

Table I: Correspondence of the numbers attributed to the genes of the Paris strain on its contigs with the numbers of SEQ ID identified in the list of sequences and position of nucleic sequences coding these genes on the sequence of these contigs with their putative function, as well as their specificity relative to the Philadelphia strain

Table II « Best-BlastP »: Putative function of the ORFs identified for the Paris strain

Table III : Correspondence of the numbers attributed to the contigs with the numbers of the SEQ ID identified in the list of sequences

Contig1	SEQ ID No. 1	Contig29	SEQ ID No. 29
Contig2	SEQ ID No. 2	Contig30	SEQ ID No. 30
Contig3	SEQ ID No. 3	Contig31	SEQ ID No. 31
Contig4	SEQ ID No. 4	Contig32	SEQ ID No. 32
Contig5	SEQ ID No. 5	Contig33	SEQ ID No. 33
Contig6	SEQ ID No. 6	Contig34	SEQ ID No. 34
Contig7	SEQ ID No. 7	Contig35	SEQ ID No. 35
Contig8	SEQ ID No. 8	Contig36	SEQ ID No. 36
Contig9	SEQ ID No. 9	Contig37	SEQ ID No. 37
Contig10	SEQ ID No. 10	Contig38	SEQ ID No. 38
Contig11	SEQ ID No. 11	Contig39	SEQ ID No. 39
Contig12	SEQ ID No. 12	Contig40	SEQ ID No. 40
Contig13	SEQ ID No. 13	Contig41	SEQ ID No. 41
Contig14	SEQ ID No. 14	Contig42	SEQ ID No. 42
Contig15	SEQ ID No. 15	Contig43	SEQ ID No. 43
Contig16	SEQ ID No. 16	Contig44	SEQ ID No. 44
Contig17	SEQ ID No. 17	Contig45	SEQ ID No. 45
Contig18	SEQ ID No. 18	Contig46	SEQ ID No. 46
Contig19	SEQ ID No. 19	Contig47	SEQ ID No. 47
Contig20	SEQ ID No. 20	Contig48	SEQ ID No. 48
Contig21	SEQ ID No. 21	Contig49	SEQ ID No. 49
Contig22	SEQ ID No. 22	Contig50	SEQ ID No. 50
Contig23	SEQ ID No. 23	Contig51	SEQ ID No. 51
Contig24	SEQ ID No. 24	Contig52	SEQ ID No. 52
Contig25	SEQ ID No. 25	Contig53	SEQ ID No. 53
Contig26	SEQ ID No. 26	Contig54	SEQ ID No. 54
Contig27	SEQ ID No. 27	Contig55	SEQ ID No. 55
Contig28	SEQ ID No. 28	Contig56	SEQ ID No. 56

Table IV: Correspondence between the contigs of Legionella pneumophila Philadelphia strain and numbering of their sequence in the list of sequences

Contig of the Philadelphia strain	**Seq Id**
Contig1	SEQ ID N°3456
Contig2	SEQ ID N°3457
Contig3	SEQ ID N°3458
Contig4	SEQ ID N°3459
Contig5	SEQ ID N°3460
Contig6	SEQ ID N°3461
Contig7	SEQ ID N°3462
Contig8	SEQ ID N°3463
Contig9	SEQ ID N°3464
Contig10	SEQ ID N°3465
Contig11	SEQ ID N°3466
Contig12	SEQ ID N°3467
Contig13	SEQ ID N°3468
Contig14	SEQ ID N°3469
Contig15	SEQ ID N°3470
Contig16	SEQ ID N°3471
Contig17	SEQ ID N°3472
Contig18	SEQ ID N°3473
Contig19	SEQ ID N°3474
Contig20	SEQ ID N°3475
Contig21	SEQ ID N°3476
Contig22	SEQ ID N°3477
Contig23	SEQ ID N°3478
Contig24	SEQ ID N°3479
Contig25	SEQ ID N°3480
Contig26	SEQ ID N°3481
Contig27	SEQ ID N°3482
Contig28	SEQ ID N°3483
Contig29	SEQ ID N°3484
Contig30	SEQ ID N°3485
Contig31	SEQ ID N°3486
Contig32	SEQ ID N°3487
Contig33	SEQ ID N°3488
Contig34	SEQ ID N°3489
Contig35	SEQ ID N°3490
Contig36	SEQ ID N°3491
Contig37	SEQ ID N°3492
Contig38	SEQ ID N°3493
Contig39	SEQ ID N°3494
Contig40	SEQ ID N°3495
Contig41	SEQ ID N°3496
Contig42	SEQ ID N°3497
Contig43	SEQ ID N°3498
Contig44	SEQ ID N°3499

Contig45	SEQ ID N°3500
Contig46	SEQ ID N°3501
Contig47	SEQ ID N°3502
Contig48	SEQ ID N°3503
Contig49	SEQ ID N°3504
Contig50	SEQ ID N°3505
Contig51	SEQ ID N°3506

Table V: Surface proteins of *Legionella pneumophila* Paris strain

The proteins of surfaces specific to the Paris strain are indicated in bold.

SEQ ID	IPF	Annotation/simiarity to other proteins
3410	**94.7**	**Surface protein, adhesion protein of *Streptococcus* sp. and *Pseudomonas* sp., Rtx toxin**
704	202.3	Surface antigen of *Bordetella* sp. and *Coxiella burnetii*
746	209.2	Surface protein of *Wolbachia* sp.
2267	440.4	Surface protein of *Mycoplasma hominis, Streptococcus*
2751	514.5	Protein similar to 440.4
3192	627.1	Surface protein of *Streptococcus pyogenes* (« collagen-like »)
3218	663.2	Lipopolysaccharide biosynthesis, O-antigen acetylase *Pseudomonas aeruginosa*
3221	667.3	Surface antigen of *Trypanosoma cruzi*
3222	668.4	IcmE of *Legionella pneumophila*
3317	803.3	Flagellar protein (« L-ring protein »)
3324	817.7	Surface protein of *Mycoplasma hominis*
136	1115.4	Rtx toxin of *Magnetococcus* sp., putative lipoprotein of *Leptospira kirschneri*
171	**1171.3**	**Protein of « surface exclusion » type of *Salmonella typhimurium***
310	1391.1	Transporter of protons of *Coxiella burnetii*
337	**1429.4**	**Rtx toxin, surface protein of *Bacillus cereus***
481	**1653.3**	**Activator of plasminogene of *Yersinia pestis*, protease associated with the cellular envelope**
527	1724.3	Hydrolase of the cellular envelope of *Pseudomonas putida*
652	**1910.6**	**Protein similar to 440.4**
664	1933.4	Protein similar to 440.4
893	2343.2	Surface antigen of *Rickettsia* sp.
972	2448.3	Hypothetical protein of *Coxiella burnetii*, periplasmic protein
1148	**2727.2**	**Surface protein of *Spirochète***
1298	2968.3	FimV, assembly of pili of *Legionella pneumophila*
1361	3059.2	Immunogene protein of *Legionella pneumophila*
1503	3271.3	O-antigen acetylase of *Pseudomnas aeruginosa*
1521	**3299.3**	**Agglutinine, adhesine, surface protein of *Brucella melitensis***
1576	3374.1	Surface antigen of *Magnetospirillum magnetotacticum*
1651	3496.2	Glycoprotein rich in histidine of *Plasmodium lophurae*
1755	3636.1	Lipoprotein *pal of Legionella pneumophila*
1847	3780.2	Surface protein of *Plasmodium falciparum*

1877	3827.2	O-antigen acetylase of *Pseudomnas aeruginosa*
2224	4347.2	Protein of the external membrane of *Burkholderia fungorum*, surface antigen of *Rickettsia* sp.
2406	4608.1	Antigen of erythrocyte infected by *Plasmodium*
2843	**5349.3**	**Major surface protein of *Anaplasma marginale*, hypothetical protein of *Plasmodium falciparum***
2930	5526.2	Adhesine, virulence protein of *Escherichia coli*
3037	**5739.3**	**Protein of « surface exclusion » type of *Pseudomonas putida*, *Enterococcus faecalis*, Rtx toxin**
3139	6037.1	Surface antigen of *Entamoeba histolytica* and *Plasmodium falciparum*
3157	6079.1	Surface antigen of *Trypanosoma cruzi*
3165	6097.1	RtxA protein of *Legionella pneumophila*
3181	**6131.1**	**Adhesine/ surface protein of *Pseudomonas putida*, *Enterococcus faecalis*, Rtx toxin**

Table VI: Proteins implied in biosynthesis of polysaccharides having a cellular envelope of *Legionella pneumophila*

SEQ ID	IPF	Annotation/similarity with other proteins
1126	269.1	Heptosyl transferase, biosynthesis of lipopolysaccharides in *Coxiella burnetii*
3218	663.2	O-acetyl transferase, modification of lipopolysaccharides in *Vibrio cholerae*
288	1360.6	Protein implied in biosynthesis of lipopolysaccharides in *Methanosarcina*
632	1882.2	Polysaccharide deacetylase of *Coxiella burnetii*
917	2371.1	Proteine CapM of *Rickettsia conorii*, glycosyltransferase
1503	3271.3	Acetylase of antigen O of *Pseudomonas aeruginosa*, modification of lipopolysaccharides
1555	3348.2	Predicted xylanase/chitine deacetylase of *Cytophaga hutchinsonii*
1877	3827.2	Acetylase of antigen O of *Pseudomonas aeruginosa*, modification of lipopolysaccharides
1928	3923.2	Potential epimerase of nucleoside-diphosphate-sugar of *Thermoanaerobacter tengcongensis*
1963	3980.1	Phosphopantetheine adenylyltransferase of *Ralstonia metallidurans*, biosynthesis of lipopolysaccharides
2204	4323.1	Pyrophosphorylase of nucleoside-diphosphate-sugar
2212	4334.1	Polysaccharide deacetylase, xylanase/chitin deacetylase
2243	4371.1	Polysaccharide deacetylase, xylanase/chitin deacetylase
2324	4488.1	Aminotransferase, synthesis of lipopolysaccharides
2378	4567.2	WciT of *Streptococcus pneumoniae*, biosynthesis of polysaccharides
2410	4616.2	Biosynthesis of antigen O, hypothetical protein of *Coxiella burnetii*
2411	4618.1	Biosynthesis of lipopolysaccharides, glycosyltransferase

5

Table X: Correspondence of the numbers attributed to the contigs of L. pneumophila philadelphia with the numbers of the SEQ ID identified in the list of sequences

Contig1	SEQ ID N∞7061
Contig2	SEQ ID N∞7062
Contig3	SEQ ID N∞7063
Contig4	SEQ ID N∞7064
Contig5	SEQ ID N∞7065
Contig6	SEQ ID N∞7066
Contig7	SEQ ID N∞7067
Contig8	SEQ ID N∞7068
Contig9	SEQ ID N∞7069
Contig10	SEQ ID N∞7070
Contig11	SEQ ID N∞7071
Contig12	SEQ ID N∞7072
Contig13	SEQ ID N∞7073

Table XI: List of the sequences of L. pneumophila philadelphia identified as specific to this strain relative to the Paris and Lens strains and position of these sequences on the contigs

Indication on the specifics of the Philadelphia strain

IPF Lp Philadelphia	Contig	SEQ ID	Position1	Position2
1007.1	CONTIG9	SEQ ID N°7069	1062411	1062962
10563.1	CONTIG13	SEQ ID N°7073	1463	2446
3775.3	CONTIG13	SEQ ID N°7073	1463	2446
1067.1	CONTIG7	SEQ ID N°7067	163133	163567
1980.3	CONTIG7	SEQ ID N°7067	163628	163918
1102.1	CONTIG7	SEQ ID N°7067	189792	190835
1109.1	CONTIG7	SEQ ID N°7067	195874	198036
4935.1	CONTIG9	SEQ ID N°7069	1604318	1605460
7686.1	CONTIG9	SEQ ID N°7069	1604318	1605460
1771.2	CONTIG8	SEQ ID N°7068	552424	553377
1773.1	CONTIG9	SEQ ID N°7069	961264	962745
1296.1	CONTIG9	SEQ ID N°7069	961264	962745
1297.1	CONTIG9	SEQ ID N°7069	959864	960817
1298.1	CONTIG9	SEQ ID N°7069	959562	959810
1302.1	CONTIG9	SEQ ID N°7069	958145	958699
1303.1	CONTIG9	SEQ ID N°7069	957452	957922
1307.1	CONTIG9	SEQ ID N°7069	956035	956523
1309.1	CONTIG9	SEQ ID N°7069	955209	955589
1310.1	CONTIG9	SEQ ID N°7069	954726	955034
1312.1	CONTIG9	SEQ ID N°7069	953857	954711
1313.1	CONTIG9	SEQ ID N°7069	953085	953864
1315.1	CONTIG9	SEQ ID N°7069	952598	953161

1319.1	CONTIG9	SEQ ID N°7069	950926	951858
1320.1	CONTIG9	SEQ ID N°7069	948772	950907
1321.1	CONTIG9	SEQ ID N°7069	948180	948743
1322.1	CONTIG9	SEQ ID N°7069	947726	948187
1323.1	CONTIG9	SEQ ID N°7069	947134	947706
1324.1	CONTIG9	SEQ ID N°7069	946011	946685
1325.1	CONTIG9	SEQ ID N°7069	945182	945862
1327.1	CONTIG9	SEQ ID N°7069	944125	945009
1328.1	CONTIG9	SEQ ID N°7069	943800	944288
1330.1	CONTIG9	SEQ ID N°7069	943233	943535
1331.1	CONTIG9	SEQ ID N°7069	942938	943243
1332.1	CONTIG9	SEQ ID N°7069	942371	942934
1333.1	CONTIG9	SEQ ID N°7069	941655	942368
1334.1	CONTIG9	SEQ ID N°7069	940367	941653
1337.1	CONTIG9	SEQ ID N°7069	940003	940242
1339.1	CONTIG9	SEQ ID N°7069	937439	940036
1340.1	CONTIG9	SEQ ID N°7069	937111	937446
1341.1	CONTIG9	SEQ ID N°7069	936506	937114
1342.1	CONTIG9	SEQ ID N°7069	935514	936509
1343.1	CONTIG9	SEQ ID N°7069	934830	935513
1344.1	CONTIG9	SEQ ID N°7069	933019	934890
1345.1	CONTIG9	SEQ ID N°7069	932244	933026
1346.1	CONTIG9	SEQ ID N°7069	931832	932254
1347.1	CONTIG9	SEQ ID N°7069	930450	931835
2328.1	CONTIG3	SEQ ID N°7063	749	1696
1348.1	CONTIG9	SEQ ID N°7069	927724	930444
1350.1	CONTIG9	SEQ ID N°7069	927256	927720
1353.1	CONTIG9	SEQ ID N°7069	925299	927188
1355.1	CONTIG9	SEQ ID N°7069	919333	925278
1356.1	CONTIG9	SEQ ID N°7069	918010	919137
1358.1	CONTIG9	SEQ ID N°7069	916581	917417
1359.1	CONTIG9	SEQ ID N°7069	916079	916570
1360.1	CONTIG9	SEQ ID N°7069	914469	916049
1361.1	CONTIG9	SEQ ID N°7069	913897	914271
1365.1	CONTIG9	SEQ ID N°7069	911529	913049
1366.1	CONTIG9	SEQ ID N°7069	910007	911458
1367.1	CONTIG9	SEQ ID N°7069	909405	909701
1368.1	CONTIG9	SEQ ID N°7069	908527	909309
1369.1	CONTIG9	SEQ ID N°7069	908004	908504
1370.1	CONTIG9	SEQ ID N°7069	907151	907993
1371.1	CONTIG9	SEQ ID N°7069	905264	907069
1373.1	CONTIG9	SEQ ID N°7069	903498	905264
1376.1	CONTIG9	SEQ ID N°7069	902217	902987
1378.1	CONTIG9	SEQ ID N°7069	900968	901987
1380.1	CONTIG9	SEQ ID N°7069	899639	900985
1382.1	CONTIG9	SEQ ID N°7069	898691	899314
1383.1	CONTIG9	SEQ ID N°7069	898266	898532
1384.1	CONTIG9	SEQ ID N°7069	897979	898281
1386.1	CONTIG9	SEQ ID N°7069	897062	898066
1435.1	CONTIG9	SEQ ID N°7069	864915	866060

1492.1	CONTIG9	SEQ ID N°7069	821283	821579
1494.1	CONTIG9	SEQ ID N°7069	820792	821109
1500.1	CONTIG9	SEQ ID N°7069	814942	815592
1501.1	CONTIG9	SEQ ID N°7069	814765	815241
1503.1	CONTIG9	SEQ ID N°7069	813201	813485
1518.1	CONTIG9	SEQ ID N°7069	800919	803852
1519.1	CONTIG9	SEQ ID N°7069	799375	800655
1520.1	CONTIG9	SEQ ID N°7069	796924	799023
1538.1	CONTIG9	SEQ ID N°7069	775795	776787
1627.1	CONTIG8	SEQ ID N°7068	454194	455027
1631.1	CONTIG8	SEQ ID N°7068	456917	457408
1635.1	CONTIG8	SEQ ID N°7068	459876	460751
1663.1	CONTIG8	SEQ ID N°7068	477274	477810
1674.1	CONTIG8	SEQ ID N°7068	488306	490213
1676.1	CONTIG8	SEQ ID N°7068	491037	491288
1723.1	CONTIG8	SEQ ID N°7068	522733	523200
1724.1	CONTIG8	SEQ ID N°7068	523306	523560
1725.1	CONTIG8	SEQ ID N°7068	523670	523945
1749.1	CONTIG8	SEQ ID N°7068	537066	537341
1760.1	CONTIG8	SEQ ID N°7068	544546	545574
1767.1	CONTIG8	SEQ ID N°7068	549553	550482
1772.1	CONTIG6	SEQ ID N°7066	779	1804
1785.1	CONTIG9	SEQ ID N°7069	970464	970808
1787.1	CONTIG9	SEQ ID N°7069	970976	971416
1892.1	CONTIG7	SEQ ID N°7067	76315	77967
3771.3	CONTIG13	SEQ ID N°7073	43581	46268
1952.1	CONTIG9	SEQ ID N°7069	1328145	1328393
1981.1	CONTIG9	SEQ ID N°7069	1347920	1348357
2005.1	CONTIG9	SEQ ID N°7069	1364695	1366614
2006.1	CONTIG9	SEQ ID N°7069	1366816	1368231
2026.1	CONTIG9	SEQ ID N°7069	1380350	1380739
2059.1	CONTIG9	SEQ ID N°7069	1407671	1408075
2066.1	CONTIG9	SEQ ID N°7069	1418764	1421481
2083.1	CONTIG9	SEQ ID N°7069	1436828	1439185
2084.1	CONTIG9	SEQ ID N°7069	1439377	1440240
2086.1	CONTIG9	SEQ ID N°7069	1440346	1441680
2125.2	CONTIG9	SEQ ID N°7069	1470612	1471472
2132.1	CONTIG11	SEQ ID N°7071	144883	145770
2133.1	CONTIG11	SEQ ID N°7071	144142	144882
2134.1	CONTIG11	SEQ ID N°7071	143221	143718
2135.1	CONTIG11	SEQ ID N°7071	143204	144145
2141.1	CONTIG11	SEQ ID N°7071	137310	137879
2202.1	CONTIG11	SEQ ID N°7071	87995	89587
2311.1	CONTIG8	SEQ ID N°7068	553457	554560
2312.1	CONTIG8	SEQ ID N°7068	554678	555862
2314.1	CONTIG8	SEQ ID N°7068	555930	556454
2315.1	CONTIG8	SEQ ID N°7068	556495	556839
2316.1	CONTIG8	SEQ ID N°7068	556651	558360
2317.1	CONTIG8	SEQ ID N°7068	558425	558796
2319.1	CONTIG8	SEQ ID N°7068	559260	560681

2321.1	CONTIG8	SEQ ID N°7068	560851	562347
2323.1	CONTIG8	SEQ ID N°7068	562541	562981
2324.1	CONTIG8	SEQ ID N°7068	563118	563813
2325.1	CONTIG8	SEQ ID N°7068	563899	564930
2326.1	CONTIG8	SEQ ID N°7068	564927	566123
2330.1	CONTIG8	SEQ ID N°7068	567385	567954
2333.1	CONTIG8	SEQ ID N°7068	568120	569727
2337.1	CONTIG8	SEQ ID N°7068	572901	573461
2343.1	CONTIG8	SEQ ID N°7068	575955	578648
2345.1	CONTIG8	SEQ ID N°7068	578635	579891
2346.1	CONTIG8	SEQ ID N°7068	579888	581024
2348.1	CONTIG8	SEQ ID N°7068	581063	584191
2350.1	CONTIG8	SEQ ID N°7068	584584	586371
2357.1	CONTIG8	SEQ ID N°7068	590170	591033
2373.1	CONTIG8	SEQ ID N°7068	601883	602209
2385.1	CONTIG8	SEQ ID N°7068	613763	615529
2465.1	CONTIG12	SEQ ID N°7072	96957	97913
2466.1	CONTIG12	SEQ ID N°7072	97888	98478
2467.1	CONTIG12	SEQ ID N°7072	98499	99803
2468.1	CONTIG12	SEQ ID N°7072	99820	100638
2469.1	CONTIG12	SEQ ID N°7072	100647	101573
2470.1	CONTIG12	SEQ ID N°7072	101679	102866
2471.1	CONTIG12	SEQ ID N°7072	103108	104379
2472.1	CONTIG12	SEQ ID N°7072	104380	105558
2473.1	CONTIG12	SEQ ID N°7072	105588	106370
2498.1	CONTIG12	SEQ ID N°7072	124135	124545
2500.1	CONTIG12	SEQ ID N°7072	123924	126110
2684.2	CONTIG8	SEQ ID N°7068	740173	740430
2710.1	CONTIG9	SEQ ID N°7069	1778548	1780281
2743.1	CONTIG9	SEQ ID N°7069	1810120	1810353
2769.1	CONTIG9	SEQ ID N°7069	1826652	1828562
2782.3	CONTIG8	SEQ ID N°7068	78469	79875
5631.2	CONTIG8	SEQ ID N°7068	78469	79875
2784.1	CONTIG8	SEQ ID N°7068	73057	73398
2889.1	CONTIG13	SEQ ID N°7073	19409	20020
2890.1	CONTIG13	SEQ ID N°7073	18305	19162
2892.1	CONTIG13	SEQ ID N°7073	17653	17907
2894.1	CONTIG13	SEQ ID N°7073	15676	17652
2895.1	CONTIG13	SEQ ID N°7073	15195	15683
2896.1	CONTIG13	SEQ ID N°7073	14143	14952
3353.1	CONTIG13	SEQ ID N°7073	14143	14952
2915.1	CONTIG13	SEQ ID N°7073	21609	22910
3909.2	CONTIG13	SEQ ID N°7073	21609	22910
2916.1	CONTIG13	SEQ ID N°7073	23001	24515
3908.2	CONTIG13	SEQ ID N°7073	23001	24515
2917.1	CONTIG13	SEQ ID N°7073	24645	24908
2918.1	CONTIG13	SEQ ID N°7073	24971	26290
2919.1	CONTIG13	SEQ ID N°7073	26453	26965
2920.1	CONTIG13	SEQ ID N°7073	27050	27442
2921.1	CONTIG13	SEQ ID N°7073	27535	28215

3018.1	CONTIG13	SEQ ID N°7073	41565	42278
10565.1	CONTIG9	SEQ ID N°7069	1320956	1322047
1947.2	CONTIG9	SEQ ID N°7069	1320956	1322047
3019.1	CONTIG9	SEQ ID N°7069	1320956	1322047
3020.1	CONTIG9	SEQ ID N°7069	1320447	1320905
3021.1	CONTIG9	SEQ ID N°7069	1319365	1320159
3023.1	CONTIG9	SEQ ID N°7069	1317520	1319235
3024.1	CONTIG9	SEQ ID N°7069	1316653	1317507
3025.1	CONTIG9	SEQ ID N°7069	1315445	1316551
3027.1	CONTIG9	SEQ ID N°7069	1314906	1315190
3046.1	CONTIG9	SEQ ID N°7069	1298926	1299282
3047.1	CONTIG9	SEQ ID N°7069	1298615	1299712
3087.1	CONTIG9	SEQ ID N°7069	1016602	1017420
3097.1	CONTIG11	SEQ ID N°7071	183622	184182
3100.1	CONTIG11	SEQ ID N°7071	182447	182740
3101.1	CONTIG11	SEQ ID N°7071	182080	182409
3102.1	CONTIG11	SEQ ID N°7071	180595	181965
3103.1	CONTIG11	SEQ ID N°7071	179123	180463
3104.1	CONTIG11	SEQ ID N°7071	177735	179093
3105.1	CONTIG11	SEQ ID N°7071	176469	177605
3317.1	CONTIG9	SEQ ID N°7069	1863795	1864346
2685.2	CONTIG13	SEQ ID N°7073	12622	13683
3356.1	CONTIG13	SEQ ID N°7073	12622	13683
3357.1	CONTIG13	SEQ ID N°7073	11421	12638
3362.1	CONTIG13	SEQ ID N°7073	7035	7835
3364.1	CONTIG13	SEQ ID N°7073	5298	6497
3365.1	CONTIG13	SEQ ID N°7073	4728	5141
34.1	CONTIG9	SEQ ID N°7069	40333	43707
35.1	CONTIG9	SEQ ID N°7069	44063	44371
3538.1	CONTIG9	SEQ ID N°7069	1883947	1885137
3539.1	CONTIG9	SEQ ID N°7069	1885249	1885500
3541.1	CONTIG9	SEQ ID N°7069	1887373	1888815
3542.1	CONTIG9	SEQ ID N°7069	1888997	1889311
3544.1	CONTIG9	SEQ ID N°7069	1889663	1889962
3545.1	CONTIG9	SEQ ID N°7069	1890063	1890284
3548.1	CONTIG9	SEQ ID N°7069	1891102	1892148
3550.2	CONTIG9	SEQ ID N°7069	1893944	1895122
3551.1	CONTIG9	SEQ ID N°7069	1895356	1895775
3552.2	CONTIG9	SEQ ID N°7069	1895768	1896094
3683.1	CONTIG8	SEQ ID N°7068	129540	131426
37.1	CONTIG9	SEQ ID N°7069	44561	44827
3740.2	CONTIG9	SEQ ID N°7069	1025799	1027007
2327.1	CONTIG12	SEQ ID N°7072	10	987
3742.1	CONTIG12	SEQ ID N°7072	10	987
1944.1	CONTIG13	SEQ ID N°7073	999	1205
3793.1	CONTIG9	SEQ ID N°7069	1491228	1491776
3794.1	CONTIG9	SEQ ID N°7069	1490472	1491212
3795.4	CONTIG9	SEQ ID N°7069	1489860	1490459
3797.2	CONTIG9	SEQ ID N°7069	1921869	1923518
3798.1	CONTIG9	SEQ ID N°7069	1923682	1924194

3799.1	CONTIG9	SEQ ID N°7069	1924210	1924524
3800.1	CONTIG9	SEQ ID N°7069	1924550	1926376
3803.1	CONTIG9	SEQ ID N°7069	1926518	1927483
3827.1	CONTIG13	SEQ ID N°7073	2622	3332
3829.2	CONTIG13	SEQ ID N°7073	3307	4557
3847.1	CONTIG9	SEQ ID N°7069	1482616	1483233
3848.1	CONTIG9	SEQ ID N°7069	1481860	1482336
3849.3	CONTIG9	SEQ ID N°7069	1480223	1481680
3850.1	CONTIG9	SEQ ID N°7069	1479788	1480216
3890.1	CONTIG13	SEQ ID N°7073	20841	21641
39.1	CONTIG9	SEQ ID N°7069	44906	45655
8044.1	CONTIG7	SEQ ID N°7067	16618	17850
3986.2	CONTIG5	SEQ ID N°7065	2	2959
4124.2	CONTIG5	SEQ ID N°7065	2	2959
3997.1	CONTIG9	SEQ ID N°7069	1920982	1921893
40.1	CONTIG9	SEQ ID N°7069	45836	46858
4018.1	CONTIG9	SEQ ID N°7069	1913464	1914630
4019.1	CONTIG9	SEQ ID N°7069	1912844	1913476
4022.2	CONTIG9	SEQ ID N°7069	1911062	1911376
4064.1	CONTIG9	SEQ ID N°7069	1486212	1486631
4065.1	CONTIG9	SEQ ID N°7069	1485967	1486812
4066.1	CONTIG9	SEQ ID N°7069	1486860	1487468
4067.4	CONTIG9	SEQ ID N°7069	1487626	1488717
41.1	CONTIG9	SEQ ID N°7069	46950	47837
43.1	CONTIG9	SEQ ID N°7069	48387	50369
4936.1	CONTIG9	SEQ ID N°7069	1605170	1605475
504.1	CONTIG9	SEQ ID N°7069	473249	474757
506.1	CONTIG9	SEQ ID N°7069	475061	476458
507.1	CONTIG9	SEQ ID N°7069	476469	477350
508.1	CONTIG9	SEQ ID N°7069	477545	477874
518.1	CONTIG9	SEQ ID N°7069	485931	487241
519.1	CONTIG9	SEQ ID N°7069	487186	487464
521.1	CONTIG9	SEQ ID N°7069	487597	487989
669.1	CONTIG9	SEQ ID N°7069	604354	605595
6736.1	CONTIG6	SEQ ID N°7066	1824	2537
6874.1	CONTIG8	SEQ ID N°7068	575999	576325
750.1	CONTIG9	SEQ ID N°7069	1271952	1274519
751.1	CONTIG9	SEQ ID N°7069	1271377	1271823
752.1	CONTIG9	SEQ ID N°7069	1270905	1271255
760.1	CONTIG9	SEQ ID N°7069	1266574	1267239
761.1	CONTIG9	SEQ ID N°7069	1266303	1266707
764.1	CONTIG9	SEQ ID N°7069	1263381	1263866
765.1	CONTIG9	SEQ ID N°7069	1261771	1262709
774.1	CONTIG9	SEQ ID N°7069	1258317	1260068
778.1	CONTIG9	SEQ ID N°7069	1256902	1257267
8067.2	CONTIG9	SEQ ID N°7069	1488930	1489379
8073.2	CONTIG9	SEQ ID N°7069	1484592	1484990
10294.1	CONTIG8	SEQ ID N°7068	119556	120962
7817.1	CONTIG8	SEQ ID N°7068	119556	120962
8134.1	CONTIG8	SEQ ID N°7068	119556	120962

874.1	CONTIG9	SEQ ID N°7069	1177144	1178064
875.1	CONTIG9	SEQ ID N°7069	1176375	1176995
876.1	CONTIG9	SEQ ID N°7069	1175483	1176361
1769.1	CONTIG3	SEQ ID N°7063	21	716
9388.1	CONTIG3	SEQ ID N°7063	21	716
4934.1	CONTIG9	SEQ ID N°7069	1819527	1820453
2798.2	CONTIG8	SEQ ID N°7068	71193	77576
3630.3	CONTIG12	SEQ ID N°7072	11054	16705

Table XII: Position of former contigs of the Paris strain on the genomic sequence of the chromosome of the Paris strain of sequence SEQ ID 3507 and of the plasmid of the Paris strain of sequence SEQ ID 3508

5

	pos1	pos2	former contig	SEQ ID of former contig
chromosome	1	44600	41	SEQ ID N∞41
	44600	161000	56	SEQ ID N∞56
	162000	223000	54	SEQ ID N∞54
	223000	232000	39	SEQ ID N∞39
	236000	424000	49	SEQ ID N∞49
	437000	469000	37	SEQ ID N∞37
	464000	758000	53	SEQ ID N∞53
	758000	781000	45	SEQ ID N∞45
	789000	879000	43	SEQ ID N∞43
	883000	901000	36	SEQ ID N∞36
	898000	986000	42	SEQ ID N∞42
	990000	1160000	46	SEQ ID N∞46
	1160000	1214000	39	SEQ ID N∞39
	1214000	1352000	54	SEQ ID N∞54
	1352000	1670000	52	SEQ ID N∞52
	1670000	1736000	40	SEQ ID N∞40
	1736000	2040000	55	SEQ ID N∞55
	2044000	2093000	56	SEQ ID N∞56
	2093000	2204722	45	SEQ ID N∞45
	2204000	2298000	50	SEQ ID N∞50
	2298000	2656000	56	SEQ ID N∞56
	2656000	2740000	50	SEQ ID N∞50
	2740000	2753600	33	SEQ ID N∞33
	2753600	2954000	47	SEQ ID N∞47
	2954000	3178000	51	SEQ ID N∞51
	3178000	3289000	44	SEQ ID N∞44
	3289000	3449000	48	SEQ ID N∞48
	3449000	3463000	34	SEQ ID N∞34
	3463000	3503610	41	SEQ ID N∞41
plasmid	1	131900	55(position1 to 132400)	SEQ ID N∞55

Table XIII: Correspondence of the numbers attributed to the chromosome and au plasmid of the Paris and Lens strain with the SEQ ID numbers identified in the list of sequences

SeqID=3507	chromosome of the Paris strain
SeqID=3508	plasmid of the Paris strain
SeqID=6733	chromosome of the Lens strain
SeqID=6734	plasmid of the Lens strain

Table XIV: Correspondence of the numbers attributed to the genes of the Paris strain on its chromosome of sequence SEQ ID 3507 and on its plasmid of sequence SEQ ID 3508 with the numbers of the SEQ ID identified in the list of sequences and position of the nucleic sequences coding these genes on the sequence of the chromosome and of the plasmid with their putative function

Table XV: Nature of the class listed in the "Class" column in Tables XIV and XVI

1. Cellular envelope and cellular processes

1.1 Cellular wall and external membrane
1.2 Proteins of transport/bond and lipoproteins
1.3 Sensors (transduction of signal)
1.4 Bioenergy of membrane
1.5 Mobility and chimiotaxia
1.6 Secretion of protein
1.7 Cellular division
1.8 Structures of cellular surface and pili

2. Intermediary metabolism

2.1 Metabolism of glucides and related molecules
2.1.1 Specific ways
2.1.2 Principal glycolytic ways
2.1.3 TCA cycle
2.2 Metabolism of aminoacids and related molecules
2.3 Metabolism of nucleotides and nucleic acids
2.4 Metabolism of lipids
2.5 Metabolism of coenzymes and prosthetic groups
2.6 Metabolism of phosphate

3. Information paths

3.1 Replication of DNA
3.2 Repair and restriction/modification of DNA

3.3 Recombination of DNA
3.4 Segregation and encapsidation of DNA
3.5 Synthesis of RNA
3.5.1 Initiation
3.5.2 Regulation
3.5.3 Elongation
3.5.4 Termination
3.6 Modification of RNA
3.7 Synthesis of protein
3.7.1 Ribosomal proteins
3.7.2 Synthetases of aminoacyl-tRNA
3.7.3 Initiation
3.7.4 Elongation
3.7.5 Termination
3.8 Modification of protein
3.9 Folding of protein

4. Other functions

4.1 Adaptation to atypical conditions
4.2 Detoxification
4.3 Toxins
4.4 Functions relating to phage
4.5 Transposon, IS, Plasmid
4.6 Various

5. Similar to unknown proteins

5.1 Of Legionella (similar though not the same)
5.2 Of other organisms

6. No similarity

Similar to the enzyme IIN of the PTS XXXX-specific system
Similar to the transcriptional regulator (family xx)
Similar to the transportor ABC (protein of ATP bond)
Similar to the transportor ABC (permease)
Similar to the transportor ABC (bond protein)
Similar to the response regulator with two compounds
Similar to the histidine kinase sensor with two compounds
Protein bound to putative peptidoglycane (LPXTG pattern)

Table XVI: Correspondence of the numbers attributed to the specific genes of the Lens strain relative to the Paris and Philadelphia strains on its chromosome of sequence SEQ ID 6733 and on its plasmid of sequence SEQ ID 6734 with the numbers of SEQ ID identified in the list of sequences and position of the nucleic sequences coding these genes on the sequence of the chromosome and the plasmid with their putative function

Table XVII: List of the specific sequences of the Paris strain relative to the Lens and Philadelphia strains with their Pasteur Institute « ORF » correspondence number and accession number in the gene banks

1056.1	SEQID=3544	EMBL_NAME=lpp1800
1067.1	SEQID=3551	EMBL_NAME=lpp1877
1069.2	SEQID=3552	EMBL_NAME=lpp1878
1076.3	SEQID=6591	EMBL_NAME=plpp0105
1077.1	SEQID=6592	EMBL_NAME=plpp0104
1078.2	SEQID=6593	EMBL_NAME=plpp0103
1080.2	SEQID=3558	EMBL_NAME=lpp3012
1081.2	SEQID=3559	EMBL_NAME=lpp3011
11.2	SEQID=3573	EMBL_NAME=lpp0196
114.2	SEQID=3598	EMBL_NAME=lpp2957
115.2	SEQID=3603	EMBL_NAME=lpp2956
116.1	SEQID=3609	EMBL_NAME=lpp2955
1160.3	SEQID=3610	EMBL_NAME=lpp0125
1171.4	SEQID=6594	EMBL_NAME=plpp0017
1172.2	SEQID=6595	EMBL_NAME=plpp0018
118.1	SEQID=3618	EMBL_NAME=lpp2954
1183.4	SEQID=3621	EMBL_NAME=lpp0356
1213.2	SEQID=3638	EMBL_NAME=lpp0257
1235.3	SEQID=6598	EMBL_NAME=plpp0121
1237.2	SEQID=6600	EMBL_NAME=plpp0119
1299.3	SEQID=6601	EMBL_NAME=plpp0036
13.1	SEQID=3688	EMBL_NAME=lpp0195
1342.3	SEQID=6602	EMBL_NAME=plpp0034
1344.4	SEQID=6603	EMBL_NAME=plpp0033
1362.3	SEQID=6604	EMBL_NAME=plpp0098
1364.2	SEQID=6605	EMBL_NAME=plpp0099
1372.2	SEQID=3726	EMBL_NAME=lpp2385
1373.1	SEQID=3727	EMBL_NAME=lpp2384
1375.2	SEQID=3728	EMBL_NAME=lpp2383
1376.2	SEQID=3729	EMBL_NAME=lpp2382
1387.2	SEQID=3735	EMBL_NAME=lpp0079
1388.2	SEQID=3736	EMBL_NAME=lpp0080
139.6	SEQID=3737	EMBL_NAME=lpp1100
1392.2	SEQID=3740	EMBL_NAME=lpp1097
1394.2	SEQID=3741	EMBL_NAME=lpp2557
1429.4	SEQID=3766	EMBL_NAME=lpp2442
1522.2	SEQID=6606	EMBL_NAME=plpp0127
1523.2	SEQID=6607	EMBL_NAME=plpp0128
1524.3	SEQID=6608	EMBL_NAME=plpp0129
1566.3	SEQID=3846	EMBL_NAME=lpp2490
1570.4	SEQID=3850	EMBL_NAME=lpp0077
1599.5	SEQID=6610	EMBL_NAME=plpp0039
1623.4	SEQID=3881	EMBL_NAME=lpp2394
1624.5	SEQID=3882	EMBL_NAME=lpp2395

163.1 SEQID=3887 EMBL_NAME=lpp2344
1655.2 SEQID=3905 EMBL_NAME=lpp1895
1683.2 SEQID=3923 EMBL_NAME=lpp0046
172.1 SEQID=3946 EMBL_NAME=lpp2978
1735.1 SEQID=3955 EMBL_NAME=lpp1862
1761.4 SEQID=6611 EMBL_NAME=plpp0074
1779.3 SEQID=3985 EMBL_NAME=lpp2040
1787.4 SEQID=3992 EMBL_NAME=lpp1824
18.1 SEQID=4000 EMBL_NAME=lpp0192
1803.2 SEQID=4003 EMBL_NAME=lpp2374
1815.2 SEQID=4010 EMBL_NAME=lpp2986
1848.5 SEQID=6613 EMBL_NAME=plpp0124
1849.4 SEQID=6614 EMBL_NAME=plpp0125
1852.2 SEQID=6615 EMBL_NAME=plpp0126
1891.2 SEQID=4056 EMBL_NAME=lpp3007
19.1 SEQID=4061 EMBL_NAME=lpp0191
1920.3 SEQID=4074 EMBL_NAME=lpp2405
1923.2 SEQID=4075 EMBL_NAME=lpp2406
1924.4 SEQID=4076 EMBL_NAME=lpp2407
2.1 SEQID=4112 EMBL_NAME=lpp0163
20.1 SEQID=4113 EMBL_NAME=lpp0190
2006.2 SEQID=6616 EMBL_NAME=plpp0032
2027.2 SEQID=4125 EMBL_NAME=lpp1910
2030.2 SEQID=4127 EMBL_NAME=lpp0241
2051.2 SEQID=6617 EMBL_NAME=plpp0035
2078.2 SEQID=4152 EMBL_NAME=lpp2456
2112.2 SEQID=4167 EMBL_NAME=lpp3006
2153.3 SEQID=4196 EMBL_NAME=lpp2427
2172.2 SEQID=6619 EMBL_NAME=plpp0107
2173.2 SEQID=6620 EMBL_NAME=plpp0106
2174.1 SEQID=6621 EMBL_NAME=plpp0102
2175.3 SEQID=6622 EMBL_NAME=plpp0101
2176.3 SEQID=6623 EMBL_NAME=plpp0139
2177.2 SEQID=6624 EMBL_NAME=plpp0140
220.1 SEQID=4215 EMBL_NAME=lpp2859
221.1 SEQID=4220 EMBL_NAME=lpp2860
2212.4 SEQID=6626 EMBL_NAME=plpp0038
222.1 SEQID=4223 EMBL_NAME=lpp2861
223.1 SEQID=4227 EMBL_NAME=lpp2862
2239.3 SEQID=4233 EMBL_NAME=lpp1330
2244.2 SEQID=6627 EMBL_NAME=plpp0052
2245.2 SEQID=6628 EMBL_NAME=plpp0053
2247.7 SEQID=6629 EMBL_NAME=plpp0054
2258.2 SEQID=6630 EMBL_NAME=plpp0130
2260.2 SEQID=6631 EMBL_NAME=plpp0131
2270.2 SEQID=4248 EMBL_NAME=lpp0082
2271.1 SEQID=4249 EMBL_NAME=lpp0081
2351.2 SEQID=4293 EMBL_NAME=lpp3016
2367.2 SEQID=4306 EMBL_NAME=lpp2386
2374.2 SEQID=4312 EMBL_NAME=lpp2987

24.1 SEQID=4325 EMBL_NAME=lpp0189
2414.2 SEQID=4338 EMBL_NAME=lpp2879
2428.2 SEQID=4349 EMBL_NAME=lpp1870
2438.2 SEQID=4356 EMBL_NAME=lpp0074
2439.2 SEQID=4357 EMBL_NAME=lpp0073
2441.2 SEQID=4359 EMBL_NAME=lpp0072
2461.2 SEQID=4373 EMBL_NAME=lpp0769
2462.2 SEQID=4374 EMBL_NAME=lpp0770
2464.2 SEQID=4375 EMBL_NAME=lpp0771
2483.2 SEQID=4389 EMBL_NAME=lpp0052
2488.2 SEQID=4392 EMBL_NAME=lpp0336
2489.2 SEQID=4393 EMBL_NAME=lpp0337
2490.2 SEQID=4395 EMBL_NAME=lpp0338
25.1 SEQID=4399 EMBL_NAME=lpp0188
254.2 SEQID=6633 EMBL_NAME=plpp0028
255.1 SEQID=6634 EMBL_NAME=plpp0029
2555.5 SEQID=4432 EMBL_NAME=lpp2147
256.1 SEQID=6635 EMBL_NAME=plpp0030
2573.3 SEQID=6636 EMBL_NAME=plpp0094
258.2 SEQID=6637 EMBL_NAME=plpp0031
26.1 SEQID=4457 EMBL_NAME=lpp0187
2605.3 SEQID=4461 EMBL_NAME=lpp0198
2616.1 SEQID=4464 EMBL_NAME=lpp0201
2622.1 SEQID=4468 EMBL_NAME=lpp0197
2625.2 SEQID=4469 EMBL_NAME=lpp2376
2626.1 SEQID=4470 EMBL_NAME=lpp2375
2659.3 SEQID=6638 EMBL_NAME=plpp0022
2660.1 SEQID=6639 EMBL_NAME=plpp0023
2661.2 SEQID=6640 EMBL_NAME=plpp0024
2662.2 SEQID=6641 EMBL_NAME=plpp0025
2665.1 SEQID=6642 EMBL_NAME=plpp0040
2666.1 SEQID=6643 EMBL_NAME=plpp0041
2667.2 SEQID=6644 EMBL_NAME=plpp0042
27.1 SEQID=4506 EMBL_NAME=lpp0186
2712.1 SEQID=4515 EMBL_NAME=lpp2620
2713.2 SEQID=4516 EMBL_NAME=lpp2621
2726.1 SEQID=4522 EMBL_NAME=lpp2636
2727.2 SEQID=4523 EMBL_NAME=lpp0256
2767.1 SEQID=4547 EMBL_NAME=lpp1561
2815.2 SEQID=4575 EMBL_NAME=lpp3070
2830.3 SEQID=4586 EMBL_NAME=lpp1928
2873.1 SEQID=6645 EMBL_NAME=plpp0109
2874.1 SEQID=6646 EMBL_NAME=plpp0108
2877.2 SEQID=6647 EMBL_NAME=plpp0100
2878.2 SEQID=6648 EMBL_NAME=plpp0097
2879.4 SEQID=6649 EMBL_NAME=plpp0096
2881.3 SEQID=6650 EMBL_NAME=plpp0095
29.1 SEQID=4611 EMBL_NAME=lpp0185
2919.1 SEQID=4624 EMBL_NAME=lpp0144
2932.3 SEQID=4633 EMBL_NAME=lpp2880

2951.2 SEQID=4643 EMBL_NAME=lpp2139
2972.3 SEQID=4663 EMBL_NAME=lpp2122
2973.2 SEQID=4664 EMBL_NAME=lpp2121
2975.2 SEQID=4665 EMBL_NAME=lpp2120
2976.1 SEQID=4666 EMBL_NAME=lpp2119
2978.1 SEQID=4667 EMBL_NAME=lpp2118
2982.1 SEQID=4668 EMBL_NAME=lpp2117
2983.1 SEQID=4669 EMBL_NAME=lpp2116
2985.1 SEQID=4670 EMBL_NAME=lpp2115
2986.1 SEQID=4671 EMBL_NAME=lpp2114
2987.1 SEQID=4672 EMBL_NAME=lpp2113
2988.1 SEQID=4673 EMBL_NAME=lpp2112
2990.1 SEQID=4674 EMBL_NAME=lpp2111
3.1 SEQID=4683 EMBL_NAME=lpp0162
30.1 SEQID=4684 EMBL_NAME=lpp0184
3005.2 SEQID=4689 EMBL_NAME=lpp2983
3009.2 SEQID=4691 EMBL_NAME=lpp2968
3011.2 SEQID=4692 EMBL_NAME=lpp0587
3016.1 SEQID=4696 EMBL_NAME=lpp0583
3017.1 SEQID=4697 EMBL_NAME=lpp0582
3023.1 SEQID=4700 EMBL_NAME=lpp0576
3030.2 SEQID=6651 EMBL_NAME=plpp0132
3031.2 SEQID=6652 EMBL_NAME=plpp0133
3035.2 SEQID=6653 EMBL_NAME=plpp0134
3036.2 SEQID=6654 EMBL_NAME=plpp0138
3101.1 SEQID=4743 EMBL_NAME=lpp0303
3109.2 SEQID=4748 EMBL_NAME=lpp0297
3110.1 SEQID=4749 EMBL_NAME=lpp0296
3113.1 SEQID=4750 EMBL_NAME=lpp0295
3114.2 SEQID=4751 EMBL_NAME=lpp0294
3116.2 SEQID=4752 EMBL_NAME=lpp0293
3134.1 SEQID=6655 EMBL_NAME=plpp0072
3135.1 SEQID=6656 EMBL_NAME=plpp0071
3136.1 SEQID=6657 EMBL_NAME=plpp0065
3137.1 SEQID=6658 EMBL_NAME=plpp0064
3138.2 SEQID=6659 EMBL_NAME=plpp0063
3231.1 SEQID=4815 EMBL_NAME=lpp2319
3232.1 SEQID=4816 EMBL_NAME=lpp2318
3233.1 SEQID=4817 EMBL_NAME=lpp2317
3234.3 SEQID=4818 EMBL_NAME=lpp2316
3257.2 SEQID=4835 EMBL_NAME=lpp2409
3258.2 SEQID=4836 EMBL_NAME=lpp2408
3299.4 SEQID=4860 EMBL_NAME=lpp2179
3338.1 SEQID=4884 EMBL_NAME=lpp2484
3343.1 SEQID=4887 EMBL_NAME=lpp2486
3348.3 SEQID=6662 EMBL_NAME=plpp0078
3349.3 SEQID=6663 EMBL_NAME=plpp0079
3350.3 SEQID=6664 EMBL_NAME=plpp0080
3351.3 SEQID=6665 EMBL_NAME=plpp0081
3352.1 SEQID=6666 EMBL_NAME=plpp0082

3353.1 SEQID=6667 EMBL_NAME=plpp0083
3354.1 SEQID=6668 EMBL_NAME=plpp0084
3355.2 SEQID=6669 EMBL_NAME=plpp0089
3356.2 SEQID=6670 EMBL_NAME=plpp0090
3357.2 SEQID=6671 EMBL_NAME=plpp0091
3394.1 SEQID=4910 EMBL_NAME=lpp1041
34.1 SEQID=4916 EMBL_NAME=lpp0182
3406.2 SEQID=4920 EMBL_NAME=lpp1050
3411.2 SEQID=4922 EMBL_NAME=lpp2059
3466.1 SEQID=4957 EMBL_NAME=lpp2864
3486.1 SEQID=4969 EMBL_NAME=lpp1110
3498.1 SEQID=6672 EMBL_NAME=plpp0019
3530.1 SEQID=4998 EMBL_NAME=lpp1823
3532.1 SEQID=4999 EMBL_NAME=lpp1822
3584.1 SEQID=5034 EMBL_NAME=lpp1098
3631.4 SEQID=6675 EMBL_NAME=plpp0058
3632.3 SEQID=6676 EMBL_NAME=plpp0051
3653.3 SEQID=5083 EMBL_NAME=lpp2412
3655.1 SEQID=5084 EMBL_NAME=lpp2413
3658.1 SEQID=5086 EMBL_NAME=lpp2420
3659.2 SEQID=5087 EMBL_NAME=lpp2424
3661.4 SEQID=5089 EMBL_NAME=lpp2425
3664.1 SEQID=5091 EMBL_NAME=lpp3018
3665.1 SEQID=5092 EMBL_NAME=lpp3019
3666.2 SEQID=5093 EMBL_NAME=lpp3020
3729.1 SEQID=5130 EMBL_NAME=lpp0204
3730.1 SEQID=5131 EMBL_NAME=lpp0205
3731.1 SEQID=5132 EMBL_NAME=lpp0206
3732.1 SEQID=5133 EMBL_NAME=lpp0207
3735.2 SEQID=5135 EMBL_NAME=lpp0209
3764.3 SEQID=5153 EMBL_NAME=lpp1868
3851.1 SEQID=5203 EMBL_NAME=lpp1492
3887.1 SEQID=5217 EMBL_NAME=lpp2439
3888.1 SEQID=5218 EMBL_NAME=lpp2438
3889.1 SEQID=5219 EMBL_NAME=lpp2437
3890.2 SEQID=5221 EMBL_NAME=lpp2436
3891.2 SEQID=5222 EMBL_NAME=lpp2435
399.2 SEQID=5282 EMBL_NAME=lpp1560
3993.2 SEQID=5284 EMBL_NAME=lpp1640
4.1 SEQID=5288 EMBL_NAME=lpp0161
400.1 SEQID=5289 EMBL_NAME=lpp1559
401.2 SEQID=5296 EMBL_NAME=lpp1558
4037.2 SEQID=5314 EMBL_NAME=lpp0065
4096.1 SEQID=5355 EMBL_NAME=lpp0772
4097.1 SEQID=5356 EMBL_NAME=lpp0773
4098.1 SEQID=5357 EMBL_NAME=lpp0774
4101.1 SEQID=5359 EMBL_NAME=lpp0775
4104.2 SEQID=5360 EMBL_NAME=lpp2390
4105.1 SEQID=5361 EMBL_NAME=lpp2389
4106.1 SEQID=5362 EMBL_NAME=lpp2388

4107.1 SEQID=5363 EMBL_NAME=lpp2387
4108.1 SEQID=5364 EMBL_NAME=lpp2381
4109.2 SEQID=5365 EMBL_NAME=lpp2380
4123.2 SEQID=5374 EMBL_NAME=lpp1077
4127.1 SEQID=5376 EMBL_NAME=lpp1075
4128.1 SEQID=5377 EMBL_NAME=lpp1074
4129.1 SEQID=5378 EMBL_NAME=lpp1073
413.5 SEQID=5379 EMBL_NAME=lpp1905
4130.3 SEQID=5380 EMBL_NAME=lpp1072
4144.3 SEQID=5393 EMBL_NAME=lpp0412
415.2 SEQID=5397 EMBL_NAME=lpp1906
4179.2 SEQID=5420 EMBL_NAME=lpp1615
4209.2 SEQID=5437 EMBL_NAME=lpp1566
4233.4 SEQID=5452 EMBL_NAME=lpp2396
4234.4 SEQID=5453 EMBL_NAME=lpp2397
4238.1 SEQID=5455 EMBL_NAME=lpp2399
4239.4 SEQID=5456 EMBL_NAME=lpp2400
4244.1 SEQID=5460 EMBL_NAME=lpp1407
4248.1 SEQID=5463 EMBL_NAME=lpp3010
425.1 SEQID=5465 EMBL_NAME=lpp2430
4251.2 SEQID=5466 EMBL_NAME=lpp3008
426.3 SEQID=5470 EMBL_NAME=lpp2428
4281.2 SEQID=5481 EMBL_NAME=lpp2047
4282.3 SEQID=5482 EMBL_NAME=lpp2046
4284.3 SEQID=5483 EMBL_NAME=lpp2045
4285.1 SEQID=5484 EMBL_NAME=lpp2044
4421.2 SEQID=5579 EMBL_NAME=lpp0211
4422.1 SEQID=5580 EMBL_NAME=lpp0212
4424.1 SEQID=5581 EMBL_NAME=lpp0213
4479.2 SEQID=5614 EMBL_NAME=lpp1573
4481.1 SEQID=6677 EMBL_NAME=plpp0050
4482.1 SEQID=6678 EMBL_NAME=plpp0049
4483.2 SEQID=6679 EMBL_NAME=plpp0048
4490.2 SEQID=5622 EMBL_NAME=lpp0068
4491.3 SEQID=5623 EMBL_NAME=lpp0069
4492.3 SEQID=5624 EMBL_NAME=lpp0070
4493.1 SEQID=5625 EMBL_NAME=lpp0071
4496.2 SEQID=5626 EMBL_NAME=lpp0075
4500.1 SEQID=5631 EMBL_NAME=lpp2085
4516.3 SEQID=5639 EMBL_NAME=lpp0078
4533.2 SEQID=5654 EMBL_NAME=lpp0779
4546.3 SEQID=5661 EMBL_NAME=lpp2967
4549.2 SEQID=6680 EMBL_NAME=plpp0118
4553.2 SEQID=5665 EMBL_NAME=lpp0291
4636.4 SEQID=5715 EMBL_NAME=lpp2168
4660.2 SEQID=5733 EMBL_NAME=lpp1956
474.2 SEQID=5787 EMBL_NAME=lpp0777
475.1 SEQID=5794 EMBL_NAME=lpp0778
4770.1 SEQID=5809 EMBL_NAME=lpp0439
4785.1 SEQID=5820 EMBL_NAME=lpp1625

4822.3 SEQID=5848 EMBL_NAME=lpp2393
485.3 SEQID=6681 EMBL_NAME=plpp0026
486.2 SEQID=6682 EMBL_NAME=plpp0027
4860.1 SEQID=5865 EMBL_NAME=lpp0983
489.1 SEQID=5886 EMBL_NAME=lpp0456
4913.2 SEQID=5896 EMBL_NAME=lpp1054
4914.3 SEQID=5897 EMBL_NAME=lpp1053
4927.1 SEQID=5905 EMBL_NAME=lpp1904
4945.2 SEQID=6683 EMBL_NAME=plpp0003
4947.2 SEQID=6684 EMBL_NAME=plpp0002
502.3 SEQID=5956 EMBL_NAME=lpp1122
5037.2 SEQID=5967 EMBL_NAME=lpp3077
505.3 SEQID=5976 EMBL_NAME=lpp2418
507.3 SEQID=5992 EMBL_NAME=lpp2416
5078.2 SEQID=6685 EMBL_NAME=plpp0004
5081.5 SEQID=6687 EMBL_NAME=plpp0005
5088.2 SEQID=6000 EMBL_NAME=lpp1576
509.1 SEQID=6001 EMBL_NAME=lpp2415
510.2 SEQID=6008 EMBL_NAME=lpp2414
5113.3 SEQID=6689 EMBL_NAME=plpp0061
5114.3 SEQID=6690 EMBL_NAME=plpp0062
5146.2 SEQID=6022 EMBL_NAME=lpp0292
5151.1 SEQID=6692 EMBL_NAME=plpp0123
5152.1 SEQID=6693 EMBL_NAME=plpp0122
5153.1 SEQID=6694 EMBL_NAME=plpp0093
5154.2 SEQID=6695 EMBL_NAME=plpp0092
52.1 SEQID=6035 EMBL_NAME=lpp0168
5201.2 SEQID=6038 EMBL_NAME=lpp1219
5202.2 SEQID=6039 EMBL_NAME=lpp1220
5204.2 SEQID=6040 EMBL_NAME=lpp1221
5208.1 SEQID=6042 EMBL_NAME=lpp2984
5224.2 SEQID=6045 EMBL_NAME=lpp1929
5242.3 SEQID=6054 EMBL_NAME=lpp2125
5288.1 SEQID=6069 EMBL_NAME=lpp0588
5297.2 SEQID=6073 EMBL_NAME=lpp2492
5307.3 SEQID=6075 EMBL_NAME=lpp1682
5321.2 SEQID=6086 EMBL_NAME=lpp1052
5322.2 SEQID=6087 EMBL_NAME=lpp1051
5387.1 SEQID=6109 EMBL_NAME=lpp0083
5390.1 SEQID=6112 EMBL_NAME=lpp0089
5474.2 SEQID=6145 EMBL_NAME=lpp2315
5517.1 SEQID=6167 EMBL_NAME=lpp2391
5554.3 SEQID=6187 EMBL_NAME=lpp1216
5555.3 SEQID=6188 EMBL_NAME=lpp1217
5576.3 SEQID=6197 EMBL_NAME=lpp2401
5605.2 SEQID=6211 EMBL_NAME=lpp0076
5621.1 SEQID=6219 EMBL_NAME=lpp2433
5623.2 SEQID=6220 EMBL_NAME=lpp2434
5626.3 SEQID=6221 EMBL_NAME=lpp2404
5634.2 SEQID=6224 EMBL_NAME=lpp0432

5660.1 SEQID=6239 EMBL_NAME=lpp1577
567.6 SEQID=6696 EMBL_NAME=plpp0037
5675.2 SEQID=6243 EMBL_NAME=lpp2123
5677.2 SEQID=6244 EMBL_NAME=lpp2124
5688.2 SEQID=6249 EMBL_NAME=lpp2587
577.2 SEQID=6270 EMBL_NAME=lpp0203
5812.1 SEQID=6278 EMBL_NAME=lpp2377
5813.1 SEQID=6279 EMBL_NAME=lpp2378
5814.1 SEQID=6280 EMBL_NAME=lpp2379
5822.1 SEQID=6282 EMBL_NAME=lpp2863
5844.1 SEQID=6697 EMBL_NAME=plpp0047
5846.2 SEQID=6698 EMBL_NAME=plpp0057
5874.2 SEQID=6289 EMBL_NAME=lpp1563
5890.1 SEQID=6292 EMBL_NAME=lpp2411
59.1 SEQID=6298 EMBL_NAME=lpp0164
591.4 SEQID=6304 EMBL_NAME=lpp2885
5987.2 SEQID=6315 EMBL_NAME=lpp2403
6015.1 SEQID=6699 EMBL_NAME=plpp0060
6016.1 SEQID=6700 EMBL_NAME=plpp0059
6029.1 SEQID=6321 EMBL_NAME=lpp0543
6052.1 SEQID=6326 EMBL_NAME=lpp1218
6071.1 SEQID=6329 EMBL_NAME=lpp0290
6072.1 SEQID=6330 EMBL_NAME=lpp0289
6117.2 SEQID=6335 EMBL_NAME=lpp0202
616.5 SEQID=6340 EMBL_NAME=lpp2884
6165.1 SEQID=6342 EMBL_NAME=lpp2245a
6217.2 SEQID=6352 EMBL_NAME=lpp1564
6224.3 SEQID=6701 EMBL_NAME=plpp0021
6227.1 SEQID=6702 EMBL_NAME=plpp0008
6229.1 SEQID=6703 EMBL_NAME=plpp0015
6230.1 SEQID=6704 EMBL_NAME=plpp0016
6231.3 SEQID=6705 EMBL_NAME=plpp0020
6263.1 SEQID=6357 EMBL_NAME=lpp0298
6273.1 SEQID=6359 EMBL_NAME=lpp3045
6314.1 SEQID=6365 EMBL_NAME=lpp1908
6429.1 SEQID=6706 EMBL_NAME=plpp0141a
659.2 SEQID=6707 EMBL_NAME=plpp0135
660.1 SEQID=6708 EMBL_NAME=plpp0136
661.3 SEQID=6709 EMBL_NAME=plpp0137
678.3 SEQID=6398 EMBL_NAME=lpp1081
680.3 SEQID=6399 EMBL_NAME=lpp1080
681.4 SEQID=6400 EMBL_NAME=lpp1079
682.3 SEQID=6401 EMBL_NAME=lpp1078
7.1 SEQID=6412 EMBL_NAME=lpp0160
713.2 SEQID=6710 EMBL_NAME=plpp0046
715.1 SEQID=6711 EMBL_NAME=plpp0045
716.1 SEQID=6712 EMBL_NAME=plpp0044
717.3 SEQID=6713 EMBL_NAME=plpp0043
737.2 SEQID=6714 EMBL_NAME=plpp0141b
739.5 SEQID=6715 EMBL_NAME=plpp0001

750.2	SEQID=6437	EMBL_NAME=lpp2613
752.2	SEQID=6438	EMBL_NAME=lpp2614
761.2	SEQID=6716	EMBL_NAME=plpp0110
762.2	SEQID=6717	EMBL_NAME=plpp0111
764.1	SEQID=6718	EMBL_NAME=plpp0112
765.1	SEQID=6719	EMBL_NAME=plpp0113
766.4	SEQID=6720	EMBL_NAME=plpp0114
783.2	SEQID=6721	EMBL_NAME=plpp0055
785.4	SEQID=6722	EMBL_NAME=plpp0056
789.4	SEQID=6456	EMBL_NAME=lpp2245b
796.2	SEQID=6723	EMBL_NAME=plpp0068
862.1	SEQID=6496	EMBL_NAME=lpp2057
863.1	SEQID=6497	EMBL_NAME=lpp2056
875.5	SEQID=6505	EMBL_NAME=lpp2314
876.1	SEQID=6506	EMBL_NAME=lpp2313
879.3	SEQID=6507	EMBL_NAME=lpp2312
884.3	SEQID=6726	EMBL_NAME=plpp0117
885.3	SEQID=6727	EMBL_NAME=plpp0116
886.4	SEQID=6728	EMBL_NAME=plpp0115
905.2	SEQID=6527	EMBL_NAME=lpp2473
946.3	SEQID=6555	EMBL_NAME=lpp2423
947.3	SEQID=6556	EMBL_NAME=lpp2422
948.3	SEQID=6557	EMBL_NAME=lpp2421
956.2	SEQID=6729	EMBL_NAME=plpp0088
957.2	SEQID=6730	EMBL_NAME=plpp0087
959.2	SEQID=6731	EMBL_NAME=plpp0086
960.2	SEQID=6732	EMBL_NAME=plpp0085
99.1	SEQID=6582	EMBL_NAME=lpp2365

Table XVIII: List of the specific sequences of the Lens strain relative to the Paris and Philadelphia strains with their « ORF » Institut Pasteur correspondence number and accession number in the gene banks

5

1001.1	SEQID=6735	EMBL_NAME=lpl1928
1002.1	SEQID=6736	EMBL_NAME=lpl1927
1003.1	SEQID=6737	EMBL_NAME=lpl1926
102.1	SEQID=6738	EMBL_NAME=lpl0183
1020.1	SEQID=6739	EMBL_NAME=lpl1915
103.1	SEQID=6740	EMBL_NAME=lpl0182
1040.1	SEQID=6742	EMBL_NAME=lpl1900
1041.1	SEQID=6743	EMBL_NAME=lpl1899
1045.1	SEQID=6745	EMBL_NAME=lpl1896
1047.1	SEQID=6746	EMBL_NAME=lpl1895
1048.1	SEQID=6747	EMBL_NAME=lpl1894
1049.2	SEQID=6748	EMBL_NAME=lpl0064
105.1	SEQID=6749	EMBL_NAME=lpl0181
1050.1	SEQID=6750	EMBL_NAME=lpl0063
1059.1	SEQID=6751	EMBL_NAME=lpl0057

106.1	SEQID=6752	EMBL_NAME=lpl0180
110.1	SEQID=6754	EMBL_NAME=lpl0177
1105.2	SEQID=6755	EMBL_NAME=lpl0019
111.1	SEQID=6756	EMBL_NAME=lpl0176
112.1	SEQID=6757	EMBL_NAME=lpl0175
113.1	SEQID=6758	EMBL_NAME=lpl0174
115.1	SEQID=6760	EMBL_NAME=lpl0173
116.1	SEQID=6761	EMBL_NAME=lpl0172
1204.1	SEQID=6763	EMBL_NAME=lpl2879
1206.1	SEQID=6764	EMBL_NAME=lpl2877
1207.1	SEQID=6765	EMBL_NAME=lpl2876
1209.1	SEQID=6766	EMBL_NAME=lpl2875
1217.1	SEQID=6767	EMBL_NAME=lpl2869
1219.1	SEQID=6768	EMBL_NAME=lpl2868
1221.1	SEQID=6769	EMBL_NAME=lpl2867
1237.2	SEQID=6770	EMBL_NAME=lpl2857
1251.1	SEQID=6773	EMBL_NAME=lpl2848
1252.1	SEQID=6774	EMBL_NAME=lpl2847
1255.2	SEQID=6775	EMBL_NAME=lpl2845
1258.1	SEQID=6776	EMBL_NAME=lpl2843
126.1	SEQID=6777	EMBL_NAME=lpl0165
127.1	SEQID=6778	EMBL_NAME=lpl0164
1274.1	SEQID=6779	EMBL_NAME=lpl2842
1275.1	SEQID=6780	EMBL_NAME=lpl2841
1276.1	SEQID=6781	EMBL_NAME=lpl2840
1278.1	SEQID=6782	EMBL_NAME=lpl2839
1279.1	SEQID=6783	EMBL_NAME=lpl2838
1280.1	SEQID=6784	EMBL_NAME=lpl2837
1283.1	SEQID=6786	EMBL_NAME=lpl2835
1284.1	SEQID=6787	EMBL_NAME=lpl2834
1285.1	SEQID=6788	EMBL_NAME=lpl2833
1296.1	SEQID=6789	EMBL_NAME=lpl2827
1297.1	SEQID=6790	EMBL_NAME=lpl2826
1321.1	SEQID=6791	EMBL_NAME=lpl2806
1422.1	SEQID=6792	EMBL_NAME=lpl2741
1535.1	SEQID=6795	EMBL_NAME=lpl0612
154.1	SEQID=6796	EMBL_NAME=lpl0146
156.1	SEQID=6797	EMBL_NAME=lpl0145
157.1	SEQID=6798	EMBL_NAME=lpl0144
159.1	SEQID=6799	EMBL_NAME=lpl0143
1697.1	SEQID=6800	EMBL_NAME=lpl0718
1718.1	SEQID=6801	EMBL_NAME=lpl0729
1824.1	SEQID=6803	EMBL_NAME=lpl0801
1826.1	SEQID=6805	EMBL_NAME=lpl0803
1827.1	SEQID=6806	EMBL_NAME=lpl0804
1834.1	SEQID=6809	EMBL_NAME=lpl0811
1955.1	SEQID=6810	EMBL_NAME=lpl0904
2110.1	SEQID=6812	EMBL_NAME=lpl1015
2137.1	SEQID=6816	EMBL_NAME=lpl1037
2141.1	SEQID=6818	EMBL_NAME=lpl1041

2142.2	SEQID=6819	EMBL_NAME=lpl1042
2145.1	SEQID=6821	EMBL_NAME=lpl1043a
2152.2	SEQID=6823	EMBL_NAME=lpl1048
2155.1	SEQID=6824	EMBL_NAME=lpl1050
2170.1	SEQID=6827	EMBL_NAME=lpl1059
2180.1	SEQID=6828	EMBL_NAME=lpl1067
2181.1	SEQID=6829	EMBL_NAME=lpl1068
2182.1	SEQID=6830	EMBL_NAME=lpl1069
2184.1	SEQID=6831	EMBL_NAME=lpl1070
2186.2	SEQID=6832	EMBL_NAME=lpl1071
2189.1	SEQID=6833	EMBL_NAME=lpl1073
2194.1	SEQID=6835	EMBL_NAME=lpl1076
2195.1	SEQID=6836	EMBL_NAME=lpl1077
2198.1	SEQID=6837	EMBL_NAME=lpl1080
2199.1	SEQID=6838	EMBL_NAME=lpl1081
2200.1	SEQID=6839	EMBL_NAME=lpl1082
2201.1	SEQID=6840	EMBL_NAME=lpl1083
2202.1	SEQID=6841	EMBL_NAME=lpl1084
2203.1	SEQID=6842	EMBL_NAME=lpl1085
2206.1	SEQID=6844	EMBL_NAME=lpl1087
2207.1	SEQID=6845	EMBL_NAME=lpl1088
2208.1	SEQID=6846	EMBL_NAME=lpl1089
2210.1	SEQID=6848	EMBL_NAME=lpl1091
2211.1	SEQID=6849	EMBL_NAME=lpl1092
2213.1	SEQID=6850	EMBL_NAME=lpl1093
2224.1	SEQID=6851	EMBL_NAME=lpl1101
2241.1	SEQID=6853	EMBL_NAME=lpl0199
2242.1	SEQID=6854	EMBL_NAME=lpl0200
2245.1	SEQID=6856	EMBL_NAME=lpl0202
2253.1	SEQID=6860	EMBL_NAME=lpl0207
2259.1	SEQID=6861	EMBL_NAME=lpl0211
2260.1	SEQID=6862	EMBL_NAME=lpl0212
2440.1	SEQID=6866	EMBL_NAME=lpl2546
2441.1	SEQID=6867	EMBL_NAME=lpl2545
2516.1	SEQID=6868	EMBL_NAME=lpl2497
2518.1	SEQID=6870	EMBL_NAME=lpl2495
2521.1	SEQID=6871	EMBL_NAME=lpl2494
2523.1	SEQID=6872	EMBL_NAME=lpl2493
2524.1	SEQID=6873	EMBL_NAME=lpl2492
2525.1	SEQID=6874	EMBL_NAME=lpl2491
2526.1	SEQID=6875	EMBL_NAME=lpl2490
2527.1	SEQID=6876	EMBL_NAME=lpl2489
2529.1	SEQID=6877	EMBL_NAME=lpl2488
2530.1	SEQID=6878	EMBL_NAME=lpl2487
2531.1	SEQID=6879	EMBL_NAME=lpl2486
2532.1	SEQID=6880	EMBL_NAME=lpl2485
2533.1	SEQID=6881	EMBL_NAME=lpl2484
2534.1	SEQID=6882	EMBL_NAME=lpl2483
2540.1	SEQID=6884	EMBL_NAME=lpl2477
2541.1	SEQID=6885	EMBL_NAME=lpl2476

2547.1	SEQID=6886	EMBL_NAME=lpl2472
2584.1	SEQID=6888	EMBL_NAME=lpl2445
2640.1	SEQID=6890	EMBL_NAME=lpl2399
2658.1	SEQID=6891	EMBL_NAME=lpl2385
266.1	SEQID=6892	EMBL_NAME=lpl0069
267.1	SEQID=6894	EMBL_NAME=lpl0068
269.1	SEQID=6895	EMBL_NAME=lpl0067
270.1	SEQID=6897	EMBL_NAME=lpl0066
2701.1	SEQID=6898	EMBL_NAME=lpl2354
2708.1	SEQID=6900	EMBL_NAME=lpl2350
2717.1	SEQID=6901	EMBL_NAME=lpl2344
2719.1	SEQID=6902	EMBL_NAME=lpl2343
272.1	SEQID=6903	EMBL_NAME=lpl0065
2720.1	SEQID=6904	EMBL_NAME=lpl2342
2722.1	SEQID=6905	EMBL_NAME=lpl2341
2723.1	SEQID=6906	EMBL_NAME=lpl2340
273.2	SEQID=6907	EMBL_NAME=lpl1893
2738.1	SEQID=6908	EMBL_NAME=lpl2330
2749.1	SEQID=6909	EMBL_NAME=lpl2323
2775.1	SEQID=6910	EMBL_NAME=lpl2309
2782.1	SEQID=6912	EMBL_NAME=lpl2305
2795.1	SEQID=6913	EMBL_NAME=lpl2295
2796.1	SEQID=6914	EMBL_NAME=lpl2294
2798.1	SEQID=6915	EMBL_NAME=lpl2293
2801.1	SEQID=6916	EMBL_NAME=lpl2292
2802.1	SEQID=6917	EMBL_NAME=lpl2291
2805.1	SEQID=6918	EMBL_NAME=lpl2289
2806.1	SEQID=6919	EMBL_NAME=lpl2288
2808.1	SEQID=6921	EMBL_NAME=lpl2286
2810.1	SEQID=6923	EMBL_NAME=lpl2284
3002.1	SEQID=6925	EMBL_NAME=lpl2148
3042.1	SEQID=6926	EMBL_NAME=lpl2114
3054.1	SEQID=6927	EMBL_NAME=lpl2107
3056.1	SEQID=6928	EMBL_NAME=lpl2106
3057.1	SEQID=6929	EMBL_NAME=lpl2105
3064.1	SEQID=6930	EMBL_NAME=lpl2100
3065.1	SEQID=6931	EMBL_NAME=lpl2099
3066.1	SEQID=6932	EMBL_NAME=lpl2098
3140.1	SEQID=6933	EMBL_NAME=lpl2049
3155.1	SEQID=6935	EMBL_NAME=lpl2038
3156.1	SEQID=6936	EMBL_NAME=lpl2037
3158.1	SEQID=6937	EMBL_NAME=lpl2036
3160.1	SEQID=6938	EMBL_NAME=lpl2035
3161.1	SEQID=6939	EMBL_NAME=lpl2034
3162.1	SEQID=6940	EMBL_NAME=lpl2033
3417.1	SEQID=6942	EMBL_NAME=lpl0216
3420.1	SEQID=6943	EMBL_NAME=lpl0217
3422.1	SEQID=6944	EMBL_NAME=lpl0218
3435.1	SEQID=6945	EMBL_NAME=lpl0226
3728.1	SEQID=6948	EMBL_NAME=lpl0552

3743.2	SEQID=6949	EMBL_NAME=lpl0565
3744.1	SEQID=6950	EMBL_NAME=lpl0566
3745.1	SEQID=6951	EMBL_NAME=lpl0567
3747.2	SEQID=6952	EMBL_NAME=lpl0568a
3748.2	SEQID=6953	EMBL_NAME=lpl0568b
3815.2	SEQID=6957	EMBL_NAME=plpl0037
3818.1	SEQID=6959	EMBL_NAME=plpl0039
3820.1	SEQID=6960	EMBL_NAME=plpl0040
3822.1	SEQID=6961	EMBL_NAME=plpl0042
3823.1	SEQID=6962	EMBL_NAME=plpl0043
3843.1	SEQID=6972	EMBL_NAME=plpl0053
3848.2	SEQID=6973	EMBL_NAME=plpl0001
3850.1	SEQID=6974	EMBL_NAME=plpl0002
3863.1	SEQID=6978	EMBL_NAME=plpl0014
3865.1	SEQID=6979	EMBL_NAME=plpl0015
3866.1	SEQID=6980	EMBL_NAME=plpl0016
3868.1	SEQID=6981	EMBL_NAME=plpl0017
3870.1	SEQID=6982	EMBL_NAME=plpl0018
3871.1	SEQID=6983	EMBL_NAME=plpl0019
3873.1	SEQID=6984	EMBL_NAME=plpl0020
3874.1	SEQID=6985	EMBL_NAME=plpl0021
3875.1	SEQID=6986	EMBL_NAME=plpl0022
3877.1	SEQID=6987	EMBL_NAME=plpl0023
3878.1	SEQID=6988	EMBL_NAME=plpl0024
3880.1	SEQID=6989	EMBL_NAME=plpl0025
3881.1	SEQID=6990	EMBL_NAME=plpl0026
3882.1	SEQID=6991	EMBL_NAME=plpl0027
3884.1	SEQID=6992	EMBL_NAME=plpl0028
3886.2	SEQID=6993	EMBL_NAME=plpl0029
3888.1	SEQID=6994	EMBL_NAME=plpl0030
3890.1	SEQID=6995	EMBL_NAME=plpl0031
3891.2	SEQID=6996	EMBL_NAME=plpl0032
3893.1	SEQID=6997	EMBL_NAME=plpl0033
3894.1	SEQID=6998	EMBL_NAME=plpl0034
3949.1	SEQID=7002	EMBL_NAME=lpl1158
3976.1	SEQID=7003	EMBL_NAME=lpl1138
3987.1	SEQID=7004	EMBL_NAME=lpl1132
4007.1	SEQID=7006	EMBL_NAME=lpl1116
4172.1	SEQID=7014	EMBL_NAME=lpl0194
4173.1	SEQID=7015	EMBL_NAME=lpl0193
4175.1	SEQID=7017	EMBL_NAME=lpl0191
4181.1	SEQID=7019	EMBL_NAME=lpl0189
4182.1	SEQID=7020	EMBL_NAME=lpl0188
4185.1	SEQID=7021	EMBL_NAME=lpl0187
4189.2	SEQID=7023	EMBL_NAME=lpl1424
4196.1	SEQID=7024	EMBL_NAME=lpl1417
4197.1	SEQID=7025	EMBL_NAME=lpl1416
4248.1	SEQID=7030	EMBL_NAME=lpl1942
4250.1	SEQID=7031	EMBL_NAME=lpl1943
4251.1	SEQID=7032	EMBL_NAME=lpl1944

4252.1	SEQID=7033	EMBL_NAME=lpl1945
4253.1	SEQID=7034	EMBL_NAME=lpl1946
4254.1	SEQID=7035	EMBL_NAME=lpl1947
560.1	SEQID=7037	EMBL_NAME=lpl1681
561.1	SEQID=7038	EMBL_NAME=lpl1680
562.1	SEQID=7039	EMBL_NAME=lpl1679
564.1	SEQID=7040	EMBL_NAME=lpl1678
688.1	SEQID=7042	EMBL_NAME=lpl1588
689.1	SEQID=7043	EMBL_NAME=lpl1587
692.1	SEQID=7045	EMBL_NAME=lpl1585
699.1	SEQID=7048	EMBL_NAME=lpl1581
700.1	SEQID=7049	EMBL_NAME=lpl1580
703.1	SEQID=7050	EMBL_NAME=lpl1579
709.1	SEQID=7053	EMBL_NAME=lpl1575
711.1	SEQID=7054	EMBL_NAME=lpl1574
760.1	SEQID=7055	EMBL_NAME=lpl1537
761.1	SEQID=7056	EMBL_NAME=lpl1536
898.1	SEQID=7057	EMBL_NAME=lpl1425
995.2	SEQID=7059	EMBL_NAME=lpl1933
997.1	SEQID=7060	EMBL_NAME=lpl1931
2144.1	SEQID=6820	EMBL_NAME=lpl1043b

Table XIX: List of the sequences present in the Paris and Lens strain though absent in the Philadelphia strain with their Pasteur Institute « ORF » correspondence number and accession number in the gene banks

1082.3	SEQID=3560	EMBL_NAME=lpp1844
1090.1	SEQID=3565	EMBL_NAME=lpp1061
1156.3	SEQID=3606	EMBL_NAME=lpp1106
119.1	SEQID=3625	EMBL_NAME=lpp2953
121.1	SEQID=3635	EMBL_NAME=lpp2952
1225.2	SEQID=6596	EMBL_NAME=plpp0013
1226.2	SEQID=6597	EMBL_NAME=plpp0012
131.2	SEQID=3694	EMBL_NAME=lpp1099
1469.4	SEQID=3786	EMBL_NAME=lpp1843
15.1	SEQID=3803	EMBL_NAME=lpp0194
1560.2	SEQID=3842	EMBL_NAME=lpp2477
16.1	SEQID=3869	EMBL_NAME=lpp0193
1737.2	SEQID=3956	EMBL_NAME=lpp1863
1875.2	SEQID=4045	EMBL_NAME=lpp2529
2.1	SEQID=4112	EMBL_NAME=lpp0163
2026.1	SEQID=4124	EMBL_NAME=lpp1909
2039.1	SEQID=4131	EMBL_NAME=lpp0243
2275.4	SEQID=6632	EMBL_NAME=plpp0014
2357.4	SEQID=4299	EMBL_NAME=lpp2054
2427.4	SEQID=4348	EMBL_NAME=lpp1869
2453.4	SEQID=4367	EMBL_NAME=lpp2450
2649.1	SEQID=4483	EMBL_NAME=lpp0667
321.3	SEQID=4804	EMBL_NAME=lpp2981
3248.1	SEQID=4829	EMBL_NAME=lpp2478
33.1	SEQID=4861	EMBL_NAME=lpp0183
3395.3	SEQID=4911	EMBL_NAME=lpp1042
3396.1	SEQID=4912	EMBL_NAME=lpp1043
3401.2	SEQID=4917	EMBL_NAME=lpp1047
341.6	SEQID=4921	EMBL_NAME=lpp0024
3413.2	SEQID=4923	EMBL_NAME=lpp2060
3414.3	SEQID=4924	EMBL_NAME=lpp2061
3499.1	SEQID=6673	EMBL_NAME=plpp0011
3500.3	SEQID=6674	EMBL_NAME=plpp0010
3563.3	SEQID=5023	EMBL_NAME=lpp0158
3594.1	SEQID=5041	EMBL_NAME=lpp0639
3600.2	SEQID=5048	EMBL_NAME=lpp2449
3601.1	SEQID=5049	EMBL_NAME=lpp2448
3657.1	SEQID=5085	EMBL_NAME=lpp2419
3734.1	SEQID=5134	EMBL_NAME=lpp0208
3744.2	SEQID=5142	EMBL_NAME=lpp1850
3763.1	SEQID=5152	EMBL_NAME=lpp1867
3871.1	SEQID=5211	EMBL_NAME=lpp2070
3872.1	SEQID=5212	EMBL_NAME=lpp2069
3878.1	SEQID=5215	EMBL_NAME=lpp2066

4039.2	SEQID=5315	EMBL_NAME=lpp0064
4040.1	SEQID=5316	EMBL_NAME=lpp0063
4045.2	SEQID=5320	EMBL_NAME=lpp0059
417.3	SEQID=5412	EMBL_NAME=lpp1907
4276.2	SEQID=5479	EMBL_NAME=lpp2048
4532.2	SEQID=5653	EMBL_NAME=lpp2049
4763.2	SEQID=5803	EMBL_NAME=lpp1088
4764.2	SEQID=5804	EMBL_NAME=lpp1087
5056.3	SEQID=5980	EMBL_NAME=lpp1578
5058.2	SEQID=5981	EMBL_NAME=lpp1579
5059.3	SEQID=5982	EMBL_NAME=lpp1580
506.3	SEQID=5983	EMBL_NAME=lpp2417
5080.6	SEQID=6686	EMBL_NAME=plpp0006
5087.2	SEQID=5999	EMBL_NAME=lpp2016
5106.3	SEQID=6688	EMBL_NAME=plpp0007
5147.1	SEQID=6691	EMBL_NAME=plpp0009
5176.1	SEQID=6026	EMBL_NAME=lpp0712
5382.1	SEQID=6107	EMBL_NAME=lpp1086
5388.1	SEQID=6110	EMBL_NAME=lpp0084
5404.2	SEQID=6117	EMBL_NAME=lpp2443
5504.4	SEQID=6163	EMBL_NAME=lpp2053
553.1	SEQID=6173	EMBL_NAME=lpp2920
5584.2	SEQID=6200	EMBL_NAME=lpp1450
5609.1	SEQID=6214	EMBL_NAME=lpp2153
58.1	SEQID=6273	EMBL_NAME=lpp0165
6036.1	SEQID=6322	EMBL_NAME=lpp1449
650.4	SEQID=6385	EMBL_NAME=lpp0668
651.3	SEQID=6386	EMBL_NAME=lpp0669
860.2	SEQID=6495	EMBL_NAME=lpp2058
9.2	SEQID=6521	EMBL_NAME=lpp0159

Table XX: List of the sequences present in the Paris and Philadelphia strain though absent in the Lens strain with their Pasteur Institute « ORF » correspondence number and accession number in the gene banks

102.1	SEQID=3519	EMBL_NAME=lpp2364
103.1	SEQID=3525	EMBL_NAME=lpp2363
104.1	SEQID=3533	EMBL_NAME=lpp2362
107.1	SEQID=3553	EMBL_NAME=lpp2361
109.1	SEQID=3564	EMBL_NAME=lpp2360
1107.2	SEQID=3578	EMBL_NAME=lpp1601
1109.3	SEQID=3579	EMBL_NAME=lpp1600
111.2	SEQID=3580	EMBL_NAME=lpp2359
1111.3	SEQID=3581	EMBL_NAME=lpp1599
1211.3	SEQID=3637	EMBL_NAME=lpp0258
1236.2	SEQID=6599	EMBL_NAME=plpp0120
1334.3	SEQID=3707	EMBL_NAME=lpp0094
1335.2	SEQID=3708	EMBL_NAME=lpp0095

1423.2	SEQID=3763	EMBL_NAME=lpp2328
1424.2	SEQID=3764	EMBL_NAME=lpp2329
1425.2	SEQID=3765	EMBL_NAME=lpp2330
1432.3	SEQID=3767	EMBL_NAME=lpp2441
152.3	SEQID=3818	EMBL_NAME=lpp2354
1548.2	SEQID=3834	EMBL_NAME=lpp2331
1549.1	SEQID=3835	EMBL_NAME=lpp2332
1550.2	SEQID=3836	EMBL_NAME=lpp2333
158.2	SEQID=3854	EMBL_NAME=lpp2341
159.1	SEQID=3862	EMBL_NAME=lpp2342
160.1	SEQID=3870	EMBL_NAME=lpp2343
1628.3	SEQID=3885	EMBL_NAME=lpp2339
1629.1	SEQID=3886	EMBL_NAME=lpp2338
1631.3	SEQID=3888	EMBL_NAME=lpp2337
1635.4	SEQID=3891	EMBL_NAME=lpp1309
1639.4	SEQID=3894	EMBL_NAME=lpp1308
168.1	SEQID=3920	EMBL_NAME=lpp2346
1682.4	SEQID=3922	EMBL_NAME=lpp0045
169.1	SEQID=3927	EMBL_NAME=lpp2347
1703.1	SEQID=3936	EMBL_NAME=lpp1942
1775.3	SEQID=3983	EMBL_NAME=lpp1130
1847.3	SEQID=4029	EMBL_NAME=lpp1089
1887.2	SEQID=4053	EMBL_NAME=lpp1940
190.3	SEQID=4062	EMBL_NAME=lpp0234
1960.2	SEQID=4096	EMBL_NAME=lpp2340
1961.3	SEQID=4097	EMBL_NAME=lpp2355
2031.2	SEQID=4128	EMBL_NAME=lpp0330
2054.2	SEQID=4138	EMBL_NAME=lpp2603
2079.2	SEQID=4153	EMBL_NAME=lpp2455
2143.5	SEQID=4186	EMBL_NAME=lpp2615
2169.2	SEQID=4205	EMBL_NAME=lpp1890
227.2	SEQID=4247	EMBL_NAME=lpp2358
228.1	SEQID=4253	EMBL_NAME=lpp2357
2544.3	SEQID=4425	EMBL_NAME=lpp2192
2591.3	SEQID=4451	EMBL_NAME=lpp2779
2637.1	SEQID=4474	EMBL_NAME=lpp2336
2639.1	SEQID=4475	EMBL_NAME=lpp2327
2646.1	SEQID=4480	EMBL_NAME=lpp0673
2730.1	SEQID=4526	EMBL_NAME=lpp0251
2808.1	SEQID=4572	EMBL_NAME=lpp0321
2849.1	SEQID=4596	EMBL_NAME=lpp1912
2938.4	SEQID=4636	EMBL_NAME=lpp2883
2991.1	SEQID=4675	EMBL_NAME=lpp2110
2992.2	SEQID=4676	EMBL_NAME=lpp2109
3163.1	SEQID=4778	EMBL_NAME=lpp0331
3190.1	SEQID=4795	EMBL_NAME=lpp1007
3191.1	SEQID=4796	EMBL_NAME=lpp1006
3205.3	SEQID=4802	EMBL_NAME=lpp2131
3207.5	SEQID=4803	EMBL_NAME=lpp2132
3250.1	SEQID=4832	EMBL_NAME=lpp2474

3508.2	SEQID=4982	EMBL_NAME=lpp2780
3588.2	SEQID=5036	EMBL_NAME=lpp1090
3663.1	SEQID=5090	EMBL_NAME=lpp3017
3699.3	SEQID=5114	EMBL_NAME=lpp1404
3701.1	SEQID=5116	EMBL_NAME=lpp1403
371.1	SEQID=5123	EMBL_NAME=lpp2039
3783.1	SEQID=5160	EMBL_NAME=lpp1144
3822.2	SEQID=5187	EMBL_NAME=lpp0851
384.5	SEQID=5197	EMBL_NAME=lpp0716
3884.2	SEQID=5216	EMBL_NAME=lpp2440
402.2	SEQID=5304	EMBL_NAME=lpp1557
4030.1	SEQID=5310	EMBL_NAME=lpp0096
4083.4	SEQID=5348	EMBL_NAME=lpp1603
4084.3	SEQID=5349	EMBL_NAME=lpp1602
4220.2	SEQID=5443	EMBL_NAME=lpp1852
423.2	SEQID=5449	EMBL_NAME=lpp2432
424.1	SEQID=5457	EMBL_NAME=lpp2431
4242.3	SEQID=5458	EMBL_NAME=lpp1405
4249.1	SEQID=5464	EMBL_NAME=lpp3009
4453.2	SEQID=5597	EMBL_NAME=lpp1447
4517.2	SEQID=5640	EMBL_NAME=lpp2497
4528.2	SEQID=5649	EMBL_NAME=lpp2052
4530.1	SEQID=5651	EMBL_NAME=lpp2051
4559.1	SEQID=5667	EMBL_NAME=lpp0287
4591.2	SEQID=5689	EMBL_NAME=lpp2508
4819.2	SEQID=5845	EMBL_NAME=lpp2498
4965.2	SEQID=5926	EMBL_NAME=lpp0829c
4966.2	SEQID=5927	EMBL_NAME=lpp0830
5060.2	SEQID=5984	EMBL_NAME=lpp0835
5289.2	SEQID=6070	EMBL_NAME=lpp0589
5328.1	SEQID=6088	EMBL_NAME=lpp1565
5337.1	SEQID=6090	EMBL_NAME=lpp2896
5340.1	SEQID=6091	EMBL_NAME=lpp1947
538.1	SEQID=6105	EMBL_NAME=lpp0859
5496.2	SEQID=6157	EMBL_NAME=lpp0829b
5656.2	SEQID=6238	EMBL_NAME=lpp2887
5723.2	SEQID=6261	EMBL_NAME=lpp1944
5871.4	SEQID=6288	EMBL_NAME=lpp2037
590.4	SEQID=6299	EMBL_NAME=lpp2886
592.3	SEQID=6305	EMBL_NAME=lpp2348
5920.3	SEQID=6306	EMBL_NAME=lpp2311
593.2	SEQID=6308	EMBL_NAME=lpp2349
594.1	SEQID=6309	EMBL_NAME=lpp2350
5999.1	SEQID=6317	EMBL_NAME=lpp0039
6002.2	SEQID=6318	EMBL_NAME=lpp0881
6110.1	SEQID=6334	EMBL_NAME=lpp0829a
615.5	SEQID=6337	EMBL_NAME=lpp1562
6151.1	SEQID=6338	EMBL_NAME=lpp0038
6159.1	SEQID=6339	EMBL_NAME=lpp2410
6160.1	SEQID=6341	EMBL_NAME=lpp2402

6178.1	SEQID=6343	EMBL_NAME=lpp0210
6180.1	SEQID=6345	EMBL_NAME=lpp1035
6186.1	SEQID=6346	EMBL_NAME=lpp0882
6195.2	SEQID=6350	EMBL_NAME=lpp0717
6285.1	SEQID=6362	EMBL_NAME=lpp2895
6309.1	SEQID=6364	EMBL_NAME=lpp1945
6318.1	SEQID=6366	EMBL_NAME=lpp1851
6320.1	SEQID=6368	EMBL_NAME=lpp1813
6322.1	SEQID=6369	EMBL_NAME=lpp1812
743.4	SEQID=6432	EMBL_NAME=lpp2368
744.4	SEQID=6433	EMBL_NAME=lpp2369
818.2	SEQID=6469	EMBL_NAME=lpp2335
819.2	SEQID=6470	EMBL_NAME=lpp2334
864.1	SEQID=6498	EMBL_NAME=lpp2055
901.2	SEQID=6525	EMBL_NAME=lpp2471
938.3	SEQID=6550	EMBL_NAME=lpp1253
96.6	SEQID=6562	EMBL_NAME=lpp2367
97.2	SEQID=6568	EMBL_NAME=lpp2366
979.3	SEQID=6574	EMBL_NAME=lpp2494
980.1	SEQID=6575	EMBL_NAME=lpp2495
981.2	SEQID=6576	EMBL_NAME=lpp2496

Table XXI: List of the sequences present in the Philadelphia and Lens strain though absent in the Paris strain with their Pasteur Institute « ORF » correspondence number and accession number in the gene banks

1038.1	SEQID=6741	EMBL_NAME=lpl1901
1043.1	SEQID=6744	EMBL_NAME=lpl1897
1073.1	SEQID=6753	EMBL_NAME=lpl0044
1130.1	SEQID=6759	EMBL_NAME=lpl2933
117.1	SEQID=6762	EMBL_NAME=lpl0171
124.1	SEQID=6771	EMBL_NAME=lpl0167
125.1	SEQID=6772	EMBL_NAME=lpl0166
1282.1	SEQID=6785	EMBL_NAME=lpl2836
1434.1	SEQID=6793	EMBL_NAME=lpl2732
148.1	SEQID=6794	EMBL_NAME=lpl0150
175.1	SEQID=6802	EMBL_NAME=lpl0132
1825.1	SEQID=6804	EMBL_NAME=lpl0802
1828.1	SEQID=6807	EMBL_NAME=lpl0805
1829.1	SEQID=6808	EMBL_NAME=lpl0806
2100.1	SEQID=6811	EMBL_NAME=lpl1006
2134.1	SEQID=6813	EMBL_NAME=lpl1034
2135.1	SEQID=6814	EMBL_NAME=lpl1035
2136.1	SEQID=6815	EMBL_NAME=lpl1036
2139.1	SEQID=6817	EMBL_NAME=lpl1039
2151.1	SEQID=6822	EMBL_NAME=lpl1047
2157.1	SEQID=6825	EMBL_NAME=lpl1051
2168.1	SEQID=6826	EMBL_NAME=lpl1058

2192.1	SEQID=6834	EMBL_NAME=lpl1075
2205.1	SEQID=6843	EMBL_NAME=lpl1086
2209.1	SEQID=6847	EMBL_NAME=lpl1090
2238.1	SEQID=6852	EMBL_NAME=lpl1110
2244.1	SEQID=6855	EMBL_NAME=lpl0201
2246.1	SEQID=6857	EMBL_NAME=lpl0203
2247.1	SEQID=6858	EMBL_NAME=lpl0204
2248.1	SEQID=6859	EMBL_NAME=lpl0205
2261.1	SEQID=6863	EMBL_NAME=lpl0213
2392.1	SEQID=6864	EMBL_NAME=lpl2580
2422.1	SEQID=6865	EMBL_NAME=lpl2558
2517.1	SEQID=6869	EMBL_NAME=lpl2496
2539.1	SEQID=6883	EMBL_NAME=lpl2478
2558.1	SEQID=6887	EMBL_NAME=lpl2465
2587.1	SEQID=6889	EMBL_NAME=lpl2443
2660.1	SEQID=6893	EMBL_NAME=lpl2384
2696.1	SEQID=6896	EMBL_NAME=lpl2358
2706.1	SEQID=6899	EMBL_NAME=lpl2351
2777.1	SEQID=6911	EMBL_NAME=lpl2308
2807.1	SEQID=6920	EMBL_NAME=lpl2287
2809.1	SEQID=6922	EMBL_NAME=lpl2285
293.1	SEQID=6924	EMBL_NAME=lpl1879
3151.1	SEQID=6934	EMBL_NAME=lpl2042
322.1	SEQID=6941	EMBL_NAME=lpl1857
3536.1	SEQID=6946	EMBL_NAME=lpl0286
3537.1	SEQID=6947	EMBL_NAME=lpl0287
3749.2	SEQID=6954	EMBL_NAME=lpl0569
3788.1	SEQID=6955	EMBL_NAME=lpl0593
3793.1	SEQID=6956	EMBL_NAME=lpl0596
3816.1	SEQID=6958	EMBL_NAME=plpl0038
3827.1	SEQID=6963	EMBL_NAME=plpl0044
3828.1	SEQID=6964	EMBL_NAME=plpl0045
3831.1	SEQID=6965	EMBL_NAME=plpl0046
3835.1	SEQID=6966	EMBL_NAME=plpl0047
3836.1	SEQID=6967	EMBL_NAME=plpl0048
3837.1	SEQID=6968	EMBL_NAME=plpl0049
3839.1	SEQID=6969	EMBL_NAME=plpl0050
3840.1	SEQID=6970	EMBL_NAME=plpl0051
3841.1	SEQID=6971	EMBL_NAME=plpl0052
3859.1	SEQID=6975	EMBL_NAME=plpl0011
3861.1	SEQID=6976	EMBL_NAME=plpl0012
3862.1	SEQID=6977	EMBL_NAME=plpl0013
3895.1	SEQID=6999	EMBL_NAME=plpl0035
3896.1	SEQID=7000	EMBL_NAME=plpl0036
3937.1	SEQID=7001	EMBL_NAME=lpl1165
3995.1	SEQID=7005	EMBL_NAME=lpl1125
4011.1	SEQID=7007	EMBL_NAME=lpl1113
4085.1	SEQID=7008	EMBL_NAME=lpl0408
4093.1	SEQID=7009	EMBL_NAME=lpl0415
4111.1	SEQID=7010	EMBL_NAME=lpl0432

4169.1	SEQID=7011	EMBL_NAME=lpl0197
4170.1	SEQID=7012	EMBL_NAME=lpl0196
4171.2	SEQID=7013	EMBL_NAME=lpl0195
4174.1	SEQID=7016	EMBL_NAME=lpl0192
4177.1	SEQID=7018	EMBL_NAME=lpl0190
4188.1	SEQID=7022	EMBL_NAME=lpl0185
4203.1	SEQID=7026	EMBL_NAME=lpl1412
4229.2	SEQID=7027	EMBL_NAME=lpl1393
4236.2	SEQID=7028	EMBL_NAME=lpl1965
4237.1	SEQID=7029	EMBL_NAME=lpl1934
442.1	SEQID=7036	EMBL_NAME=lpl1768
566.1	SEQID=7041	EMBL_NAME=lpl1676
690.1	SEQID=7044	EMBL_NAME=lpl1586
694.1	SEQID=7046	EMBL_NAME=lpl1584
697.1	SEQID=7047	EMBL_NAME=lpl1582
705.1	SEQID=7051	EMBL_NAME=lpl1578
707.1	SEQID=7052	EMBL_NAME=lpl1577
899.2	SEQID=7058	EMBL_NAME=lpl2032

SEQ ID No.	ORF	Contig	Sens	Position 1	Position 2	Signal P	Specific to Paris/ Philadelphia	% homology retained ORFs	Presence in L. longbeachae	Legionella specific
SEQ ID No. 3098	5906.1	Contig29	m	2	541	-	+			-
SEQ ID No. 3097	5905.1	Contig29	p	747	1703	-	-	97%		-
SEQ ID No. 3181	6131.1	Contig31	p	1	2313		+			-
SEQ ID No. 3180	6130.1	Contig32	m	74	1048	-	+			-
SEQ ID No. 3179	6128.1	Contig32	m	1045	2223	-	+			-
SEQ ID No. 3178	6125.1	Contig32	m	2799	3074	-	+			-
SEQ ID No. 3177	6123.1	Contig32	m	3005	3130		+			+
SEQ ID No. 3176	6122.1	Contig32	p	3158	3457		+			+
SEQ ID No. 3175	6121.1	Contig33	p	1	570		+			-
SEQ ID No. 1494	3257.2	Contig33	p	933	1391	-	+			-
SEQ ID No. 1495	3258.2	Contig33	p	1363	2307	-	+			-
SEQ ID No. 659	1924.4	Contig33	m	2329	2898	-	+			-
SEQ ID No. 658	1923.2	Contig33	m	2892	3113	-	+			+
SEQ ID No. 657	1920.3	Contig33	m	3738	4808		+			-
SEQ ID No. 2987	5626.2	Contig33	m	4971	5936	-	+			-
SEQ ID No. 3122	5987.1	Contig33	m	5866	6501		+			+
SEQ ID No. 3089	5898.2	Contig34	m	2	862	+	+			+
SEQ ID No. 2792	523.2	Contig34	m	1137	2051	+	-	94%		-
SEQ ID No. 2795	524.2	Contig34	p	2016	3128	-	-	98%		-
SEQ ID No. 2804	526.2	Contig34	p	3100	3735	-	-	97%		-
SEQ ID No. 2812	528.2	Contig34	m	3861	4526	-	-	99%		-
SEQ ID No. 2816	529.2	Contig34	m	4519	4866	-	-	100%		-
SEQ ID No. 1736	3609.1	Contig34	p	5143	5745	-	-	99%		-
SEQ ID No. 1737	3610.1	Contig34	p	5784	6974	-	-	99%	+	+
SEQ ID No. 1738	3611.2	Contig34	m	7103	7888	-	-	98%		-
SEQ ID No. 1739	3612.2	Contig34	m	8026	8814		-	96%		-
SEQ ID No. 1740	3613.3	Contig34	p	9024	10349	-	-	97%		-
SEQ ID No. 1651	3496.2	Contig36	p	96	590		-	100%		-
SEQ ID No. 1650	3494.2	Contig36	p	706	1026	+	-	100%		-
SEQ ID No. 789	216.2	Contig36	p	1033	2307	-	-	100%		-

SEQ ID No. 782	215.1	Contig36	p	2289	2777	-	-	100%		-
SEQ ID No. 773	214.1	Contig36	p	2823	3278	-	-	100%		-
SEQ ID No. 765	213.1	Contig36	p	3226	4227	-	-	100%		-
SEQ ID No. 759	212.1	Contig36	p	4308	4799	-	-	100%		-
SEQ ID No. 755	211.1	Contig36	p	4667	5851	-	-	100%		-
SEQ ID No. 746	209.2	Contig36	m	5843	6589	+	-	100%		-
SEQ ID No. 2930	5526.2	Contig36	m	6590	7255	+	-	100%		-
SEQ ID No. 1037	2549.3	Contig36	p	7391	8842	-	-	99%		+
SEQ ID No. 1530	331.2	Contig36	m	8894	10513	-	-	100%		-
SEQ ID No. 1542	333.3	Contig36	p	10752	14054	-	-	100%		-
SEQ ID No. 3021	5690.2	Contig36	p	14241	14747	+	-	100%		-
SEQ ID No. 2651	4972.3	Contig36	p	14883	15563	-	-	100%		-
SEQ ID No. 2649	4969.3	Contig36	p	15689	16312	-	+			-
SEQ ID No. 2648	4968.2	Contig36	p	16308	16703	-	-	100%		-
SEQ ID No. 2597	4883.2	Contig37	p	189	647	-	-	100%		-
SEQ ID No. 2598	4884.1	Contig37	p	629	1198	-	-	99%		-
SEQ ID No. 2599	4885.2	Contig37	p	1299	1742	-	-	100%		-
SEQ ID No. 2968	5592.2	Contig37	p	1743	2447	-	-	100%		-
SEQ ID No. 753	2106.2	Contig37	p	2608	3153	-	-	100%		-
SEQ ID No. 752	2104.1	Contig37	p	3175	3564	-	-	100%		-
SEQ ID No. 130	1106.5	Contig37	p	3626	7762	-	-	99%		-
SEQ ID No. 489	1663.6	Contig37	p	7788	12056	-	-	99%		-
SEQ ID No. 2971	5596.3	Contig37	p	12164	12550	-	-	100%		-
SEQ ID No. 2970	5595.2	Contig37	p	12544	13098	-	-	100%		-
SEQ ID No. 2969	5593.1	Contig37	m	12853	13182		-	98%		-
SEQ ID No. 2305	4468.2	Contig37	p	13110	15197	-	-	99%	+	-
SEQ ID No. 2304	4466.3	Contig37	p	15176	16408	-	-	100%	+	-
SEQ ID No. 2303	4465.3	Contig37	m	15697	16179		-	100%		-
SEQ ID No. 2302	4464.1	Contig37	p	16396	16731	-	-	100%		-
SEQ ID No. 2301	4463.2	Contig37	p	16754	17416	-	-	99%		-
SEQ ID No. 2300	4462.3	Contig37	p	17413	18021	-	-	99%		-
SEQ ID No. 3084	5892.1	Contig37	p	18006	18314	-	-	100%		-
SEQ ID No. 3085	5894.3	Contig37	p	18311	19153	-	-	100%		-
SEQ ID No. 3086	5895.3	Contig37	p	19163	19450	-	-	100%		-
SEQ ID No. 187	1199.4	Contig37	p	19451	19795	-	-	100%		-
SEQ ID No. 188	1201.1	Contig37	p	19786	20454	-	-	100%		-

SEQ ID No. 189	1202.2	Contig37	p	20459	20884	-	-	100%		-
SEQ ID No. 2073	4133.1	Contig37	p	20866	21078	-	+			-
SEQ ID No. 2074	4134.1	Contig37	p	21059	21334	-	-	100%		-
SEQ ID No. 2075	4135.1	Contig37	p	21381	21788	-	-	100%		-
SEQ ID No. 2076	4136.1	Contig37	p	21789	22130	-	-	99%		-
SEQ ID No. 2077	4138.1	Contig37	p	22137	22697	-	-	100%		-
SEQ ID No. 2078	4139.1	Contig37	p	22698	23012	-	-	100%		-
SEQ ID No. 2079	4140.1	Contig37	p	22960	23430	-	-	99%		-
SEQ ID No. 2080	4141.1	Contig37	p	23421	23987	-	-	100%		-
SEQ ID No. 2081	4142.3	Contig37	p	23941	24357	-	-	100%		-
SEQ ID No. 2082	4143.3	Contig37	p	24358	24873	-	-	100%		-
SEQ ID No. 2083	4144.3	Contig37	p	24855	25061	-	+			-
SEQ ID No. 2084	4145.3	Contig37	p	25049	25495	-	-	100%		-
SEQ ID No. 2085	4148.3	Contig37	p	25483	26826	-	-	99%		-
SEQ ID No. 2086	4149.1	Contig37	p	27019	27390	-	-	100%		-
SEQ ID No. 2088	4151.1	Contig37	p	27411	27812	-	-	100%		-
SEQ ID No. 2089	4153.1	Contig37	p	27826	28449	-	+			-
SEQ ID No. 2090	4154.2	Contig37	p	28468	29460	-	-	99%		-
SEQ ID No. 2941	5540.1	Contig37	p	29461	29862	-	-	99%	+	-
SEQ ID No. 2940	5539.1	Contig37	m	29484	29870	-	-	98%		+
SEQ ID No. 2939	5538.2	Contig37	m	29929	30444	-	-	99%		-
SEQ ID No. 3087	5896.1	Contig37	m	30492	30977	-	+			-
SEQ ID No. 3088	5897.1	Contig37	p	30844	31131	-	+			-
SEQ ID No. 3173	6119.1	Contig39	m	1	531	-	+			-
SEQ ID No. 2988	5630.1	Contig39	p	577	1083		-	91%		-
SEQ ID No. 3048	5774.2	Contig39	p	1095	2222		-	79%		+
SEQ ID No. 741	2080.3	Contig39	p	2669	3373	-	-	94%		-
SEQ ID No. 1955	397.2	Contig39	p	3491	4015	-	-	100%	+	-
SEQ ID No. 1962	398.2	Contig39	p	3990	5279	-	-	96%	+	-
SEQ ID No. 1586	3394.1	Contig39	p	6091	6318	-	+			+
SEQ ID No. 1587	3395.2	Contig39	m	6445	7326		+			+
SEQ ID No. 1588	3396.1	Contig39	p	7740	8249		+			+
SEQ ID No. 1589	3397.1	Contig39	m	8412	8972	-	-	89%		-
SEQ ID No. 1590	3398.1	Contig39	m	8984	9424	-	-	87%		-
SEQ ID No. 1591	3399.1	Contig39	m	9506	9724	-	-	92%		+
SEQ ID No. 3106	5935.1	Contig39	p	9719	9898		-	87%		+

SEQ ID No. 1593	3401.2	Contig39	p	10196	10510	-	+		+
SEQ ID No. 1594	3402.2	Contig39	m	10565	11563	-	-	95%	-
SEQ ID No. 1595	3404.2	Contig39	m	11457	12302	-	-	91%	-
SEQ ID No. 1596	3406.2	Contig39	p	12883	13545		+		+
SEQ ID No. 2832	5322.1	Contig39	m	13558	13917	-	+		-
SEQ ID No. 2831	5321.1	Contig39	m	13921	14205	-	+		-
SEQ ID No. 2615	4914.2	Contig39	p	14302	14595	+	+		+
SEQ ID No. 2614	4913.2	Contig39	m	14931	15887		+		+
SEQ ID No. 3226	677.6	Contig39	m	16069	18072	+	-	95%	-
SEQ ID No. 3225	675.6	Contig39	m	18045	18512		-	89%	+
SEQ ID No. 3224	673.3	Contig39	m	18568	20133	-	-	97%	+
SEQ ID No. 1712	3575.2	Contig39	m	20124	21500	+	-	98%	-
SEQ ID No. 1711	3574.1	Contig39	m	21501	22601	-	-	95%	-
SEQ ID No. 116	1089.2	Contig39	m	22468	25227	-	-	96%	-
SEQ ID No. 118	1090.1	Contig39	m	25228	25578		-	87%	-
SEQ ID No. 119	1091.2	Contig39	m	25532	26836	+	-	95%	-
SEQ ID No. 120	1092.3	Contig39	m	26788	27693	-	-	98%	-
SEQ ID No. 2684	5031.2	Contig39	m	27551	28276	-	-	97%	+
SEQ ID No. 2685	5032.1	Contig39	m	28200	28625	-	-	98%	+
SEQ ID No. 2686	5033.1	Contig39	m	28574	28948	+	-	99%	+
SEQ ID No. 2829	5319.1	Contig39	m	29202	29537	-	-	97%	+
SEQ ID No. 2828	5318.2	Contig39	m	29538	30179	-	-	96%	-
SEQ ID No. 2072	4132.3	Contig39	m	30161	30625	-	-	99%	-
SEQ ID No. 2071	4131.3	Contig39	m	30592	31416	+	-	94%	-
SEQ ID No. 3049	5776.1	Contig39	m	31406	31666	-	+		+
SEQ ID No. 2070	4130.2	Contig39	m	31603	32019	+	+		-
SEQ ID No. 2068	4129.1	Contig39	m	31995	32513	+	+		-
SEQ ID No. 2067	4128.1	Contig39	m	32401	32640	-	+		-
SEQ ID No. 2066	4127.1	Contig39	m	32613	33146	-	+		+
SEQ ID No. 2065	4126.2	Contig39	m	32977	33888	-	+		+
SEQ ID No. 2064	4123.2	Contig39	p	34046	34300		-	75%	-
SEQ ID No. 3230	682.2	Contig39	p	34304	35275	-	+		-
SEQ ID No. 3229	681.3	Contig39	m	35343	36095	-	+		+
SEQ ID No. 3228	680.3	Contig39	m	36188	36706	-	+		-
SEQ ID No. 3227	678.3	Contig39	m	36741	37847		+		+
SEQ ID No. 2639	4955.4	Contig39	m	37811	39142	-	-	85%	-

SEQ ID No. 2638	4954.2	Contig39	m	39111	39632	-	-	93%	-
SEQ ID No. 2637	4952.1	Contig39	m	39610	40287	-	-	81%	+
SEQ ID No. 2636	4951.2	Contig39	m	40214	40783	-	-	90%	-
SEQ ID No. 2858	5382.1	Contig39	p	40834	41157	-	+		+
SEQ ID No. 2515	4764.2	Contig39	p	41153	41398	-	+		+
SEQ ID No. 2514	4763.2	Contig39	m	41667	42947	-	+		-
SEQ ID No. 609	1847.3	Contig39	p	43268	45091	-	-	97%	-
SEQ ID No. 1718	3588.2	Contig39	p	45192	45809	-	-	98%	-
SEQ ID No. 1717	3586.2	Contig39	p	46123	47250	-	-	88%	-
SEQ ID No. 207	1229.3	Contig39	p	47244	47789	-	-	95%	+
SEQ ID No. 206	1227.1	Contig39	p	48300	49013	-	-	95%	-
SEQ ID No. 926	2383.4	Contig39	p	49143	49373	+	+		-
SEQ ID No. 309	1390.5	Contig39	p	49497	49979	-	-	96%	+
SEQ ID No. 310	1391.1	Contig39	p	50090	50704		-	97%	-
SEQ ID No. 311	1392.2	Contig39	m	50780	51103	-	+		-
SEQ ID No. 1716	3584.1	Contig39	m	51193	51399	-	+		-
SEQ ID No. 260	131.2	Contig39	m	51668	52141		+		-
SEQ ID No. 3172	6117.1	Contig39	p	52705	55920	-	+		-
SEQ ID No. 3047	577.2	Contig39	p	56130	57014	-	+		-
SEQ ID No. 1816	3729.1	Contig39	m	57019	57618	-	+		-
SEQ ID No. 1817	3730.1	Contig39	m	57909	58367		+		+
SEQ ID No. 1818	3731.1	Contig39	m	58368	58904	-	+		-
SEQ ID No. 1819	3732.1	Contig39	p	59289	60173		+		-
SEQ ID No. 1820	3734.1	Contig39	p	60441	61568		+		-
SEQ ID No. 1821	3735.2	Contig39	p	61850	63517	-	+		+
SEQ ID No. 3079	5874.1	Contig40	m	2	553	-	+		-
SEQ ID No. 3185	615.5	Contig40	p	534	1337	-	+		-
SEQ ID No. 1176	2767.1	Contig40	m	2006	3424	-	+		+
SEQ ID No. 1969	399.2	Contig40	m	3481	5943		+		+
SEQ ID No. 1978	400.1	Contig40	m	5934	6746	-	+		-
SEQ ID No. 1985	401.2	Contig40	p	6795	8066	-	+		-
SEQ ID No. 1993	402.2	Contig40	p	8321	9913	+	+		+
SEQ ID No. 760	2120.2	Contig40	p	10617	11687	-	-	97%	+
SEQ ID No. 107	1075.3	Contig40	p	11840	13171	-	-	99%	-
SEQ ID No. 106	1074.3	Contig40	p	13158	15182	-	-	98%	-
SEQ ID No. 761	2121.2	Contig40	p	15227	15976	+	-	97%	-

SEQ ID No. 1175	2765.2	Contig40	m	16013	16903		-	95%		+
SEQ ID No. 1174	2763.2	Contig40	m	16914	17111	-	+			+
SEQ ID No. 1173	2761.1	Contig40	p	17411	17758	-	-	100%		-
SEQ ID No. 1172	2760.1	Contig40	p	17730	17999	-	+			-
SEQ ID No. 1170	2759.1	Contig40	p	17938	18912	-	-	98%		-
SEQ ID No. 1169	2757.1	Contig40	p	18913	19374	-	-	98%		-
SEQ ID No. 1168	2755.1	Contig40	p	19564	21600	-	-	98%		+
SEQ ID No. 1167	2754.2	Contig40	m	21696	22277	-	-	99%		-
SEQ ID No. 1166	2753.1	Contig40	m	22278	22781		-	98%		+
SEQ ID No. 1165	2752.1	Contig40	m	22782	23303	+	-	100%		-
SEQ ID No. 1164	2751.1	Contig40	m	23593	24363		-	99%		-
SEQ ID No. 3351	859.2	Contig40	p	24703	26091	-	-	99%	+	-
SEQ ID No. 3350	858.2	Contig40	p	26045	26566	-	-	97%		-
SEQ ID No. 371	1493.6	Contig40	p	26618	31495	-	-	99%	+	-
SEQ ID No. 509	1697.2	Contig40	p	31501	32064	-	-	98%		-
SEQ ID No. 510	1698.2	Contig40	m	32095	33180		-	96%		-
SEQ ID No. 2882	5429.1	Contig40	p	33195	33449	-	-	97%		-
SEQ ID No. 1047	2562.2	Contig40	p	33385	33834		-	95%		+
SEQ ID No. 1046	2560.3	Contig40	p	33902	34498	-	-	99%		-
SEQ ID No. 1044	2559.3	Contig40	p	34609	35655	-	-	99%		-
SEQ ID No. 2881	5428.1	Contig40	p	35656	36054	-	-	100%		-
SEQ ID No. 2880	5427.1	Contig40	p	36055	36756	-	-	100%		-
SEQ ID No. 2879	5426.2	Contig40	p	36750	37220	-	-	100%		-
SEQ ID No. 1010	2504.4	Contig40	p	37208	38215	-	-	99%		-
SEQ ID No. 1009	2501.2	Contig40	p	38209	39480	+	-	99%		-
SEQ ID No. 1572	3367.1	Contig40	p	39438	40316		-	98%		-
SEQ ID No. 3241	699.3	Contig40	m	40345	41484	-	-	96%		-
SEQ ID No. 3243	701.2	Contig40	m	41465	42997	-	-	96%		-
SEQ ID No. 3244	702.2	Contig40	m	42928	43755	-	-	96%		-
SEQ ID No. 1571	3365.2	Contig40	m	43915	44994	-	-	95%		-
SEQ ID No. 1570	3363.2	Contig40	m	44988	45944	-	-	99%		-
SEQ ID No. 1569	3362.1	Contig40	p	46177	46929	-	-	97%		-
SEQ ID No. 1568	3360.1	Contig40	p	46986	47717	+	-	89%		+
SEQ ID No. 184	1190.4	Contig40	m	47773	49917	+	-	96%		-
SEQ ID No. 1566	3359.2	Contig40	m	50090	52309		-	99%		-
SEQ ID No. 3026	5701.2	Contig40	m	52481	53623	-	-	99%		-

SEQ ID No. 2799	5250.2	Contig40	m	53596	54588	-	-	100%		-
SEQ ID No. 1029	2535.2	Contig40	m	54563	55675	-	-	100%		-
SEQ ID No. 1030	2536.2	Contig40	p	55808	57139	-	-	95%		-
SEQ ID No. 2349	4523.1	Contig40	p	57063	57629	-	-	96%		-
SEQ ID No. 2350	4524.1	Contig40	p	57787	58506	+	-	91%		-
SEQ ID No. 2351	4525.2	Contig40	m	58554	60278	-	-	98%		-
SEQ ID No. 2730	5094.2	Contig40	p	60305	61030	-	-	98%		-
SEQ ID No. 2729	5093.3	Contig40	p	61021	61755	-	-	98%		-
SEQ ID No. 2728	5092.3	Contig40	p	61746	62498	-	-	98%		+
SEQ ID No. 2800	5253.1	Contig40	m	62587	63249	-	-	99%		-
SEQ ID No. 2472	4705.3	Contig40	p	63281	64168	-	-	98%		-
SEQ ID No. 2473	4706.3	Contig40	p	64231	64728	-	-	97%		-
SEQ ID No. 2474	4709.1	Contig40	p	64725	65897	-	-	99%		-
SEQ ID No. 2475	4710.1	Contig40	p	66072	66860	+	-	98%		-
SEQ ID No. 2476	4711.1	Contig40	p	66832	67404	-	-	98%		-
SEQ ID No. 2477	4712.2	Contig40	p	67379	68674	-	-	97%		-
SEQ ID No. 2801	5254.1	Contig40	p	68650	69366	-	-	99%		-
SEQ ID No. 2802	5255.1	Contig40	p	69367	70545	+	-	99%		-
SEQ ID No. 2658	4987.2	Contig40	p	70541	71938	-	-	99%	+	-
SEQ ID No. 2584	4865.2	Contig40	m	72015	72449	-	-	100%		-
SEQ ID No. 3074	5857.1	Contig40	m	72492	72932	-	-	97%		-
SEQ ID No. 3073	5856.1	Contig40	m	72914	73102		+			-
SEQ ID No. 3080	5875.1	Contig41	m	1	303		+			-
SEQ ID No. 1882	3837.3	Contig41	p	521	2047	-	-	92%		-
SEQ ID No. 1881	3835.2	Contig41	p	2208	2951		-	96%		-
SEQ ID No. 1051	2569.3	Contig41	p	3162	3887	-	-	95%		-
SEQ ID No. 1050	2568.1	Contig41	m	3950	4378		-	98%		+
SEQ ID No. 1049	2567.2	Contig41	m	4563	5330	-	-	99%		-
SEQ ID No. 1880	3834.1	Contig41	p	5573	6613	-	-	96%		+
SEQ ID No. 1879	3832.1	Contig41	p	6654	7868		-	97%		-
SEQ ID No. 1878	3830.2	Contig41	p	7864	8817	-	-	96%		+
SEQ ID No. 1549	334.5	Contig41	m	8929	11910	-	-	97%		-
SEQ ID No. 1557	335.1	Contig41	p	12010	12723	-	-	98%		-
SEQ ID No. 1567	336.1	Contig41	p	12734	13735	-	-	99%		-
SEQ ID No. 1574	337.2	Contig41	m	13745	15148	-	-	92%	+	-
SEQ ID No. 1583	339.4	Contig41	p	15296	15880	+	+			-

SEQ ID No. 1597	341.6	Contig41	p	16002	16445	+	+		+
SEQ ID No. 2856	5380.3	Contig41	p	16574	17140	-	-	97%	-
SEQ ID No. 2263	4395.2	Contig41	m	17210	19366	-	-	99%	-
SEQ ID No. 2264	4396.1	Contig41	m	19330	19821	-	-	98%	-
SEQ ID No. 2265	4398.2	Contig41	m	20000	21013	-	-	99%	-
SEQ ID No. 2996	5644.3	Contig41	m	21132	22817	+	-	98%	-
SEQ ID No. 2619	4919.1	Contig41	m	23043	24431	-	-	99%	-
SEQ ID No. 2618	4918.2	Contig41	m	24412	25359	-	-	99%	-
SEQ ID No. 2995	5642.1	Contig41	m	25302	26252	-	-	99%	+
SEQ ID No. 1911	3895.2	Contig41	m	26209	27321	+	-	99%	-
SEQ ID No. 1910	3892.3	Contig41	m	27384	28613	-	-	99%	-
SEQ ID No. 2784	5216.2	Contig41	m	29032	29880	-	-	97%	-
SEQ ID No. 3375	891.3	Contig41	p	29993	31594	-	-	98%	+
SEQ ID No. 3374	889.1	Contig41	p	31734	32231	+	-	99%	-
SEQ ID No. 3373	888.3	Contig41	m	32336	33607	-	-	99%	-
SEQ ID No. 3444	986.3	Contig41	m	33608	33868	-	-	100%	-
SEQ ID No. 3445	989.3	Contig41	p	34035	35312	-	-	98%	+
SEQ ID No. 3447	991.2	Contig41	p	35832	36698	-	+		-
SEQ ID No. 3400	925.3	Contig41	p	36685	38610	-	-	99%	-
SEQ ID No. 84	1040.3	Contig41	p	38592	39653	-	-	96%	-
SEQ ID No. 85	1041.3	Contig41	m	40023	42467	-	-	99%	-
SEQ ID No. 939	2406.3	Contig41	m	42737	43828	-	-	97%	-
SEQ ID No. 938	2404.2	Contig41	m	43795	44931	-	-	98%	-
SEQ ID No. 937	2402.3	Contig41	m	44912	46285	-	-	100%	-
SEQ ID No. 2687	5037.1	Contig41	p	46971	47156	-	+		-
SEQ ID No. 2688	5038.1	Contig41	p	47157	47501	-	-	100%	+
SEQ ID No. 2689	5039.2	Contig41	p	47465	47713	-	-	100%	-
SEQ ID No. 2690	5040.2	Contig41	p	47714	49393	+	-	99%	-
SEQ ID No. 173	1174.2	Contig41	p	49394	50734	-	-	98%	-
SEQ ID No. 174	1176.5	Contig41	p	51021	52868	-	-	98%	+
SEQ ID No. 1048	2565.4	Contig41	m	52865	53674	-	-	97%	-
SEQ ID No. 1205	2815.2	Contig41	m	53870	55639	-	+		-
SEQ ID No. 1206	2816.2	Contig41	p	56182	58359	+	-	99%	-
SEQ ID No. 1207	2818.1	Contig41	m	58410	59606	-	-	98%	-
SEQ ID No. 1209	2820.1	Contig41	m	59607	60461	-	-	98%	-
SEQ ID No. 1210	2822.1	Contig41	p	60447	62309	+	-	97%	-

SEQ ID No. 763	2127.2	Contig41	p	62316	62687	-	-	99%		-
SEQ ID No. 764	2129.1	Contig41	p	62699	63313	-	-	98%		-
SEQ ID No. 766	2130.1	Contig41	p	63314	63889	+	-	98%		-
SEQ ID No. 767	2132.2	Contig41	p	63867	64478		-	98%		-
SEQ ID No. 768	2133.1	Contig41	m	64548	64796	-	-	98%		+
SEQ ID No. 769	2135.1	Contig41	p	64932	65342	+	-	99%		-
SEQ ID No. 770	2136.1	Contig41	p	65284	66171	-	-	100%		-
SEQ ID No. 771	2137.1	Contig41	p	66209	66487	-	-	100%		-
SEQ ID No. 772	2138.2	Contig41	p	66491	67009	-	-	100%		-
SEQ ID No. 774	2140.2	Contig41	p	66999	67553	-	-	98%		-
SEQ ID No. 412	1552.3	Contig41	p	67559	69124	-	-	100%		-
SEQ ID No. 413	1555.2	Contig41	p	69211	70080	-	-	99%		-
SEQ ID No. 414	1556.2	Contig41	p	70081	71469	-	-	100%		-
SEQ ID No. 1211	2824.1	Contig41	p	71460	71903	-	-	100%		-
SEQ ID No. 3389	910.4	Contig41	p	72529	74787	-	-	98%	+	-
SEQ ID No. 3388	909.2	Contig41	m	74894	75811	-	-	98%		-
SEQ ID No. 758	2119.2	Contig41	p	75964	77661	+	-	98%		-
SEQ ID No. 1212	2826.1	Contig41	p	78109	79644	-	-	99%		+
SEQ ID No. 1213	2827.3	Contig41	p	79904	82522	-	-	99%		+
SEQ ID No. 1214	2828.3	Contig41	p	82606	82911	-	+			-
SEQ ID No. 2927	5521.1	Contig42	p	297	830	-	-	100%		-
SEQ ID No. 2926	5520.1	Contig42	p	831	1391	+	+			+
SEQ ID No. 1793	3684.2	Contig42	p	1367	2128	-	-	100%		+
SEQ ID No. 1794	3685.1	Contig42	p	2210	2536		-	100%		-
SEQ ID No. 1795	3686.1	Contig42	p	2533	4083	-	-	100%		-
SEQ ID No. 1796	3687.1	Contig42	p	4196	5224	-	-	100%		-
SEQ ID No. 1797	3688.1	Contig42	m	5536	5817	-	-	100%		-
SEQ ID No. 3298	776.3	Contig42	m	6700	8211	-	-	100%		-
SEQ ID No. 3299	778.2	Contig42	m	8247	9074	-	-	100%		+
SEQ ID No. 3300	779.4	Contig42	p	9150	10607	-	-	99%		-
SEQ ID No. 2062	4119.2	Contig42	p	10583	11368	-	-	100%		-
SEQ ID No. 2061	4118.3	Contig42	p	11344	12006	-	-	100%		-
SEQ ID No. 2060	4116.3	Contig42	p	12007	12717	-	-	100%		-
SEQ ID No. 2059	4115.1	Contig42	p	12639	13736	-	-	100%		-
SEQ ID No. 555	1764.3	Contig42	p	13730	14869	-	-	100%		-
SEQ ID No. 556	1765.4	Contig42	p	14850	15467	-	-	100%		-

SEQ ID No. 592	1818.4	Contig42	p	15448	16980	-	-	100%	-
SEQ ID No. 2058	4113.1	Contig42	p	17043	17633	-	-	100%	-
SEQ ID No. 2057	4112.1	Contig42	p	17634	18518	-	-	100%	-
SEQ ID No. 2056	4111.2	Contig42	p	18460	19557	-	-	100%	-
SEQ ID No. 2920	5500.1	Contig42	p	19572	21095	-	-	100%	-
SEQ ID No. 2918	5499.1	Contig42	p	21199	22125	-	-	100%	-
SEQ ID No. 2917	5498.1	Contig42	p	22116	23102	-	-	100%	-
SEQ ID No. 2916	5497.1	Contig42	p	23074	24090	+	-	100%	-
SEQ ID No. 3171	6110.1	Contig42	p	24187	24447		-	100%	+
SEQ ID No. 2915	5496.2	Contig42	p	24434	24973		-	100%	+
SEQ ID No. 2646	4965.2	Contig42	p	24949	25446		-	100%	+
SEQ ID No. 2647	4966.2	Contig42	m	25547	27379	-	-	100%	+
SEQ ID No. 666	1935.4	Contig42	m	27346	28494	-	-	100%	+
SEQ ID No. 665	1934.4	Contig42	m	28475	29512	-	-	100%	+
SEQ ID No. 2530	4782.3	Contig42	m	29466	30560	-	-	100%	-
SEQ ID No. 2932	5528.2	Contig42	m	30570	31292	-	-	100%	+
SEQ ID No. 2705	5060.2	Contig42	p	31581	33122	-	-	100%	+
SEQ ID No. 2706	5061.4	Contig42	p	33331	35037	-	-	100%	+
SEQ ID No. 2091	4156.3	Contig42	p	34991	35854	-	-	100%	-
SEQ ID No. 2092	4157.2	Contig42	p	35845	37275	-	-	100%	-
SEQ ID No. 2093	4158.1	Contig42	p	37269	38429	+	-	100%	+
SEQ ID No. 2094	4159.1	Contig42	m	38443	39345		-	100%	-
SEQ ID No. 287	1359.2	Contig42	p	40114	41187	-	-	100%	-
SEQ ID No. 2095	4160.3	Contig42	m	41300	42163		-	100%	-
SEQ ID No. 2096	4161.3	Contig42	m	42188	43207	-	-	100%	-
SEQ ID No. 2934	5530.2	Contig42	p	43383	44126	-	-	100%	-
SEQ ID No. 2803	5256.2	Contig42	p	44544	44879	-	-	100%	-
SEQ ID No. 1877	3827.2	Contig42	m	45005	46999	-	-	100%	-
SEQ ID No. 1876	3826.1	Contig42	m	47150	48148	-	-	100%	-
SEQ ID No. 977	2457.2	Contig42	m	48149	48919	-	-	100%	-
SEQ ID No. 976	2456.2	Contig42	p	48965	49921	-	-	100%	-
SEQ ID No. 1875	3824.2	Contig42	p	49884	51209	-	-	100%	-
SEQ ID No. 1874	3822.2	Contig42	p	51526	52095	-	-	100%	-
SEQ ID No. 1873	3821.1	Contig42	p	52422	53600	-	-	100%	+
SEQ ID No. 880	2320.3	Contig42	p	53588	55039	-	-	99%	-
SEQ ID No. 879	2319.3	Contig42	m	55138	56562	-	-	100%	-

SEQ ID No. 941	2409.3	Contig42	p	56686	57420	+	-	100%		-
SEQ ID No. 943	2410.2	Contig42	p	57574	58845	+	-	100%		-
SEQ ID No. 1196	2795.1	Contig42	p	58861	59223	-	-	100%		-
SEQ ID No. 2866	540.3	Contig42	p	59184	59567	-	-	100%		-
SEQ ID No. 2862	539.3	Contig42	p	59488	60189	-	-	100%		-
SEQ ID No. 2855	538.1	Contig42	p	60700	61350	-	-	99%		+
SEQ ID No. 2850	537.2	Contig42	p	61594	62016	+	-	100%		+
SEQ ID No. 1082	262.3	Contig42	m	62121	63479	-	-	100%		-
SEQ ID No. 1089	263.1	Contig42	m	63539	65284	-	-	100%	+	-
SEQ ID No. 1093	264.2	Contig42	m	65185	66576	-	-	100%	+	-
SEQ ID No. 3331	829.2	Contig42	p	66690	68411	+	-	100%		-
SEQ ID No. 1195	2794.2	Contig42	m	68614	70335	-	-	100%		-
SEQ ID No. 1194	2793.1	Contig42	m	70547	71560		-	100%		-
SEQ ID No. 1193	2792.1	Contig42	p	71391	71660	-	+			+
SEQ ID No. 1192	2791.1	Contig42	p	71734	74157	-	-	100%		-
SEQ ID No. 1191	2788.1	Contig42	p	74393	75838		-	100%		-
SEQ ID No. 1190	2785.1	Contig42	p	75829	76677	-	-	100%		-
SEQ ID No. 623	1865.3	Contig42	p	76665	77765	-	-	100%		-
SEQ ID No. 622	1864.3	Contig42	p	77689	79155	-	-	100%		-
SEQ ID No. 1189	2784.3	Contig42	p	79238	79564	+	-	100%		-
SEQ ID No. 2569	4841.2	Contig42	p	79724	80797	-	-	100%		-
SEQ ID No. 2568	4837.1	Contig42	p	80754	81683	-	-	100%		-
SEQ ID No. 2567	4836.1	Contig42	p	81668	82159		-	100%		-
SEQ ID No. 2566	4835.4	Contig42	m	82095	82778	-	-	100%		-
SEQ ID No. 2565	4834.4	Contig42	p	82282	82476		-	100%		+
SEQ ID No. 3001	5650.3	Contig42	p	82949	83656	-	-	100%		-
SEQ ID No. 3000	5649.1	Contig42	p	83688	85022	-	-	100%		-
SEQ ID No. 2999	5648.1	Contig42	m	84306	84641		-	100%		-
SEQ ID No. 2998	5647.1	Contig42	p	85391	85759	-	-	100%		-
SEQ ID No. 2997	5645.2	Contig42	p	85779	88058	-	-	100%		-
SEQ ID No. 3170	6108.1	Contig42	p	88359	88874	-	+			-
SEQ ID No. 342	1440.2	Contig43	p	2	1249	+	-	100%		-
SEQ ID No. 70	1021.3	Contig43	p	1363	4413	-	-	100%		-
SEQ ID No. 1519	3295.2	Contig43	p	4645	7437	-	-	100%		-
SEQ ID No. 531	1733.3	Contig43	p	7541	8806	+	-	100%		-
SEQ ID No. 1518	3291.2	Contig43	p	8871	9356		-	100%		+

SEQ ID No. 1517	3290.1	Contig43	m	9457	10470	-	-	99%		+
SEQ ID No. 1516	3288.3	Contig43	p	10583	11272	-	-	97%		-
SEQ ID No. 1515	3285.3	Contig43	p	11250	12326	-	-	95%		-
SEQ ID No. 2358	4533.2	Contig43	m	12519	18197	+	+			-
SEQ ID No. 2505	475.1	Contig43	p	18407	18700		+			+
SEQ ID No. 2498	474.2	Contig43	p	18701	19033		+			+
SEQ ID No. 2049	4101.1	Contig43	p	19176	19379		+			-
SEQ ID No. 2047	4098.1	Contig43	p	19655	20044	-	+			-
SEQ ID No. 2046	4097.1	Contig43	p	20045	20302	-	+			-
SEQ ID No. 2045	4096.1	Contig43	m	20278	20760		+			-
SEQ ID No. 983	2464.2	Contig43	m	20673	21026		+			-
SEQ ID No. 982	2462.2	Contig43	p	21047	21406	-	+			-
SEQ ID No. 981	2461.2	Contig43	p	21548	21760		+			-
SEQ ID No. 2809	5273.1	Contig43	m	21791	22153	-	+			-
SEQ ID No. 980	2460.3	Contig43	m	22069	23802	+	-	99%		-
SEQ ID No. 1463	3212.2	Contig43	m	23799	24497	+	-	98%		-
SEQ ID No. 74	1028.3	Contig43	m	24521	25792	-	-	97%		-
SEQ ID No. 76	1030.3	Contig43	m	25732	26601	-	-	96%		-
SEQ ID No. 77	1032.3	Contig43	p	26819	27790	-	-	99%		-
SEQ ID No. 78	1034.3	Contig43	p	27791	28777	+	-	96%		+
SEQ ID No. 762	2122.2	Contig43	p	28807	30006	+	-	98%		-
SEQ ID No. 293	1367.3	Contig43	m	30046	30639	+	-	97%		-
SEQ ID No. 295	1369.2	Contig43	m	30471	30911	-	-	97%		-
SEQ ID No. 294	1368.2	Contig43	p	30475	30666		-	96%		+
SEQ ID No. 296	1370.5	Contig43	m	31018	31629	-	-	99%		-
SEQ ID No. 1071	2599.4	Contig43	m	31671	33359	-	-	99%	+	-
SEQ ID No. 1073	2600.4	Contig43	m	33387	34418	-	-	99%	+	-
SEQ ID No. 1464	3217.3	Contig43	m	34496	35749	-	-	96%		-
SEQ ID No. 1465	3218.1	Contig43	p	35810	36511	-	+			+
SEQ ID No. 1467	3221.1	Contig43	p	36457	37890	-	-	100%		-
SEQ ID No. 1468	3223.1	Contig43	p	37875	38759	-	-	98%	+	-
SEQ ID No. 1469	3224.3	Contig43	p	38946	41255	+	-	100%	+	-
SEQ ID No. 491	1666.3	Contig43	p	41339	43054	-	-	96%		+
SEQ ID No. 1348	304.2	Contig43	m	43061	45910	-	-	95%		+
SEQ ID No. 1324	300.2	Contig43	p	46000	47781	-	-	98%		-
SEQ ID No. 593	1819.6	Contig43	m	48139	52452	-	-	95%		+

SEQ ID No. 974	2450.3	Contig43	m	52511	54343	-	-	98%		-
SEQ ID No. 340	1438.4	Contig43	p	54447	56333	-	-	98%		-
SEQ ID No. 2783	521.2	Contig43	m	56667	58832	-	-	98%		-
SEQ ID No. 2776	520.1	Contig43	m	59041	59676	-	-	99%	+	-
SEQ ID No. 2771	519.3	Contig43	m	59604	61277	-	-	100%		-
SEQ ID No. 1235	2869.1	Contig43	m	61278	61583	-	-	100%		-
SEQ ID No. 1234	2868.1	Contig43	p	61643	63514	+	-	98%		-
SEQ ID No. 1233	2866.1	Contig43	p	63598	64923	-	-	97%	+	-
SEQ ID No. 1052	2570.2	Contig43	m	64993	66153	+	-	99%		-
SEQ ID No. 1053	2571.2	Contig43	m	66126	67073	-	-	97%		-
SEQ ID No. 1232	2863.1	Contig43	m	67045	67788	-	-	94%		-
SEQ ID No. 1231	2862.1	Contig43	m	67685	68956	-	-	96%		-
SEQ ID No. 156	1148.2	Contig43	m	69161	71929	-	-	98%		-
SEQ ID No. 541	1746.2	Contig43	p	71971	72792	-	-	91%		-
SEQ ID No. 542	1747.2	Contig43	m	73047	73409	-	-	100%		+
SEQ ID No. 1230	2861.1	Contig43	p	73575	74438	-	-	97%		-
SEQ ID No. 969	2443.2	Contig43	p	74407	75246	-	-	95%		+
SEQ ID No. 968	2442.2	Contig43	m	75266	76591	-	-	99%		-
SEQ ID No. 2502	4746.1	Contig43	m	76572	76862	+	-	97%		-
SEQ ID No. 2503	4748.1	Contig43	m	76881	77648	-	-	99%		-
SEQ ID No. 2504	4749.2	Contig43	m	77687	79966	-	-	98%		-
SEQ ID No. 1442	3172.2	Contig43	m	80019	80918	-	-	99%		-
SEQ ID No. 3066	5824.1	Contig43	m	80785	81201	-	+			-
SEQ ID No. 1443	3173.1	Contig43	p	81385	81654	+	-	98%		-
SEQ ID No. 1444	3174.1	Contig43	p	81648	82529	-	-	96%		-
SEQ ID No. 1445	3175.1	Contig43	p	82516	83301	-	-	94%		-
SEQ ID No. 1446	3177.1	Contig43	p	83277	83762	-	-	98%		-
SEQ ID No. 1447	3178.1	Contig43	p	83973	84629	-	-	98%		-
SEQ ID No. 1448	3179.1	Contig43	p	84611	86452	+	-	99%		-
SEQ ID No. 3169	6107.1	Contig43	m	86295	86489		+			+
SEQ ID No. 261	1314.3	Contig43	m	86777	88240	+	-	99%		-
SEQ ID No. 1450	3181.2	Contig43	m	88225	89379	-	-	99%		-
SEQ ID No. 2762	5167.4	Contig43	m	89396	89701	-	+			-
SEQ ID No. 3168	6106.1	Contig43	m	89715	89972		+			-
SEQ ID No. 1275	2938.3	Contig44	p	205	3858		-	98%		+
SEQ ID No. 1274	2934.1	Contig44	m	3925	5202	-	-	97%		+

SEQ ID No. 1273	2933.2	Contig44	m	5328	6305	-	-	94%		-
SEQ ID No. 1272	2932.2	Contig44	m	6328	8805	-	+			-
SEQ ID No. 2833	5324.1	Contig44	m	8763	9419		+			-
SEQ ID No. 946	2414.2	Contig44	m	9382	10485	-	+			-
SEQ ID No. 945	2413.1	Contig44	m	10652	10867	-	+			-
SEQ ID No. 944	2412.2	Contig44	m	11014	12684	-	-	97%		-
SEQ ID No. 1074	2602.2	Contig44	m	12685	13578	-	-	100%		-
SEQ ID No. 1075	2604.3	Contig44	m	13579	15414	-	-	99%	+	-
SEQ ID No. 3281	755.3	Contig44	m	15517	16344	-	-	97%		-
SEQ ID No. 3280	754.2	Contig44	m	16437	17450	-	-	98%		-
SEQ ID No. 3279	753.2	Contig44	p	17613	18584	+	-	98%		-
SEQ ID No. 1270	2929.1	Contig44	p	18836	19348	-	-	99%		-
SEQ ID No. 1269	2928.2	Contig44	m	19362	20231	-	-	99%		-
SEQ ID No. 1268	2926.1	Contig44	p	20406	20564		+			+
SEQ ID No. 1267	2925.1	Contig44	p	20611	22620	-	-	99%		-
SEQ ID No. 1266	2924.3	Contig44	m	22704	24785		-	96%		+
SEQ ID No. 3065	5823.2	Contig44	p	24801	26072	+	-	99%		-
SEQ ID No. 845	226.4	Contig44	m	26153	27034	-	-	96%		+
SEQ ID No. 1634	3466.1	Contig44	p	27076	27321	-	+			-
SEQ ID No. 3064	5822.1	Contig44	p	27230	27472	-	+			+
SEQ ID No. 826	223.1	Contig44	m	27631	28530	-	+			-
SEQ ID No. 822	222.1	Contig44	p	28556	29101	-	+			-
SEQ ID No. 817	221.1	Contig44	p	29082	29993	-	+			-
SEQ ID No. 812	220.1	Contig44	p	29972	30685	-	+			-
SEQ ID No. 649	1906.2	Contig44	m	31080	32615	-	-	99%		-
SEQ ID No. 648	1905.2	Contig44	m	32600	33892	-	-	99%		-
SEQ ID No. 349	1454.2	Contig44	m	34019	35995	-	-	100%		-
SEQ ID No. 2846	536.3	Contig44	p	36111	38714	-	-	98%		-
SEQ ID No. 2835	533.2	Contig44	p	38923	39735	-	-	96%		-
SEQ ID No. 2830	532.2	Contig44	m	39785	40900	+	-	99%		-
SEQ ID No. 784	2151.2	Contig44	p	41026	42417	-	-	96%		-
SEQ ID No. 1456	3196.3	Contig44	p	42479	43987	-	-	99%	+	-
SEQ ID No. 884	2324.4	Contig44	p	44152	45291	-	-	98%		+
SEQ ID No. 3063	5821.1	Contig44	p	45218	45577	-	+			+
SEQ ID No. 883	2323.2	Contig44	m	45462	46322	-	-	97%		+
SEQ ID No. 882	2322.2	Contig44	m	46545	47288	-	-	98%		-

SEQ ID No. 1457	3198.1	Contig44	m	47234	47797	-	-	100%		-
SEQ ID No. 1458	3199.1	Contig44	m	48054	49808	-	-	97%		-
SEQ ID No. 3218	663.2	Contig44	p	49960	51945	-	-	98%		-
SEQ ID No. 3219	665.2	Contig44	m	51996	52349	-	-	98%		-
SEQ ID No. 3220	666.2	Contig44	p	52242	52985	-	-	100%		-
SEQ ID No. 1459	3201.1	Contig44	p	53078	55018	+	-	100%		-
SEQ ID No. 3196	632.2	Contig44	p	54982	55926	-	-	98%		-
SEQ ID No. 3197	633.3	Contig44	p	55984	57357	-	-	100%		-
SEQ ID No. 2805	5266.2	Contig44	m	57427	60807	-	-	89%		+
SEQ ID No. 3029	5719.2	Contig44	p	60988	61779	-	-	98%		-
SEQ ID No. 1300	297.2	Contig44	p	61710	62090	-	-	99%		-
SEQ ID No. 1289	296.1	Contig44	p	62252	62626	+	-	99%		-
SEQ ID No. 1281	295.1	Contig44	p	62864	63349	-	-	100%		-
SEQ ID No. 1276	294.1	Contig44	p	63372	64055	-	-	99%	+	-
SEQ ID No. 1271	293.2	Contig44	p	64030	65301	-	-	98%	+	-
SEQ ID No. 1263	292.1	Contig44	p	65323	65865	-	-	99%	+	-
SEQ ID No. 1255	291.4	Contig44	p	65866	67161	-	-	99%		-
SEQ ID No. 66	1011.5	Contig44	p	67175	69529	-	-	98%		-
SEQ ID No. 65	1010.2	Contig44	p	69530	70564	-	-	97%		-
SEQ ID No. 64	1009.2	Contig44	p	70565	71083	-	-	98%		-
SEQ ID No. 63	1006.3	Contig44	p	71092	71754	+	-	96%		-
SEQ ID No. 1742	3617.2	Contig44	p	71658	72077	-	-	100%		-
SEQ ID No. 1743	3618.2	Contig44	p	72070	74055	+	-	99%		-
SEQ ID No. 318	1402.2	Contig44	p	74062	75576	-	-	100%		-
SEQ ID No. 317	1400.2	Contig44	p	75570	77072	-	-	99%		-
SEQ ID No. 1744	3619.1	Contig44	p	77256	77771	-	-	100%		-
SEQ ID No. 1745	3621.2	Contig44	p	77720	79261	-	-	99%		-
SEQ ID No. 530	1730.3	Contig44	p	79326	81956	-	-	99%		-
SEQ ID No. 1927	3920.1	Contig44	p	81957	82331	-	-	100%		-
SEQ ID No. 1926	3919.1	Contig44	p	82297	83229	-	-	98%		-
SEQ ID No. 1925	3917.1	Contig44	p	83374	83652	-	+			-
SEQ ID No. 3113	5957.1	Contig44	p	83695	84627		+			-
SEQ ID No. 1924	3916.3	Contig44	p	84579	85955		-	97%		-
SEQ ID No. 1017	2517.2	Contig44	p	86081	87181	+	-	95%		-
SEQ ID No. 1018	2518.4	Contig44	m	87240	87869	-	-	99%		-
SEQ ID No. 1020	2520.4	Contig44	m	87791	88135	-	-	98%		-

SEQ ID No. 2691	5041.2	Contig44	p	88262	88798	-	-	100%		-
SEQ ID No. 2692	5042.1	Contig44	m	88860	89207		-	98%		+
SEQ ID No. 2693	5043.3	Contig44	p	88939	90213	+	-	98%		-
SEQ ID No. 2708	5064.2	Contig44	p	90483	91244	-	-	97%		+
SEQ ID No. 2707	5062.3	Contig44	p	91473	92150	-	-	94%		-
SEQ ID No. 3062	5818.1	Contig44	m	91982	92296		+			+
SEQ ID No. 2142	4232.3	Contig44	m	92543	93277	-	-	97%		-
SEQ ID No. 2141	4231.2	Contig44	p	93455	94969	-	-	97%		+
SEQ ID No. 3399	923.3	Contig44	m	95217	97235		-	97%		+
SEQ ID No. 3398	922.2	Contig44	m	97211	98074	+	-	98%		-
SEQ ID No. 3397	921.3	Contig44	m	98064	98678	-	-	98%		-
SEQ ID No. 3061	5816.1	Contig44	m	98666	99019	-	-	99%		-
SEQ ID No. 3338	84.4	Contig44	m	99038	100735	-	-	96%		-
SEQ ID No. 3327	82.2	Contig44	m	101521	101775		-	100%	+	-
SEQ ID No. 1157	2741.2	Contig44	m	102053	102259	-	+			-
SEQ ID No. 1158	2742.2	Contig44	p	102309	102515	-	+			-
SEQ ID No. 3305	79.1	Contig44	p	102657	102878	-	+			-
SEQ ID No. 3292	77.1	Contig44	p	103230	103448	-	+			+
SEQ ID No. 3031	5720.1	Contig45	p	1	204		+			+
SEQ ID No. 3032	5723.1	Contig45	p	218	523	-	+			-
SEQ ID No. 3033	5724.2	Contig45	p	651	2195	-	-	89%		+
SEQ ID No. 513	1703.1	Contig45	m	2232	3020	-	-	98%		-
SEQ ID No. 514	1704.4	Contig45	m	3189	5213	-	-	97%		-
SEQ ID No. 636	1887.2	Contig45	m	5447	7078		-	90%		-
SEQ ID No. 637	1889.2	Contig45	m	7392	8750	-	-	91%		+
SEQ ID No. 1614	3432.2	Contig45	p	8936	10051	-	-	95%		-
SEQ ID No. 1613	3431.3	Contig45	p	10009	11346	-	-	98%		-
SEQ ID No. 839	2250.2	Contig45	p	11577	12230	+	-	95%		+
SEQ ID No. 1611	3429.1	Contig45	p	12616	14904	-	-	95%		+
SEQ ID No. 1610	3428.1	Contig45	m	15087	15905	-	-	96%		-
SEQ ID No. 1609	3426.1	Contig45	m	16078	16782		-	97%		-
SEQ ID No. 1608	3423.3	Contig45	m	17254	18540	-	-	91%		-
SEQ ID No. 1607	3421.2	Contig45	p	18675	20015	-	-	96%		+
SEQ ID No. 1605	3419.3	Contig45	m	20115	20882	-	-	92%		+
SEQ ID No. 2788	5224.2	Contig45	p	21155	21526	-	+			+
SEQ ID No. 1216	2830.3	Contig45	p	21481	22104	-	+			+

SEQ ID No. 1217	2831.1	Contig45	p	22331	23041	-	-	82%	-
SEQ ID No. 1218	2833.1	Contig45	p	23134	24129		-	91%	-
SEQ ID No. 508	1695.2	Contig45	m	24311	25426	-	-	92%	-
SEQ ID No. 507	1693.2	Contig45	p	25663	25947	-	-	91%	-
SEQ ID No. 1219	2835.1	Contig45	p	26014	26910	+	-	91%	-
SEQ ID No. 1220	2838.1	Contig45	p	26911	28107	-	-	92%	-
SEQ ID No. 1222	2840.1	Contig45	p	28121	29647		-	91%	-
SEQ ID No. 1223	2841.4	Contig45	p	29814	31133	-	-	95%	-
SEQ ID No. 288	1360.6	Contig45	m	31446	32963	-	-	96%	-
SEQ ID No. 289	1361.3	Contig45	p	33558	34631	-	-	98%	-
SEQ ID No. 246	1286.3	Contig45	p	34628	36538	-	-	98%	-
SEQ ID No. 245	1284.2	Contig45	p	36772	37059	-	-	96%	-
SEQ ID No. 244	1283.2	Contig45	p	37016	38341	-	-	98%	-
SEQ ID No. 1224	2847.4	Contig45	p	38461	39105	-	-	95%	+
SEQ ID No. 1225	2848.4	Contig45	m	39163	39426		+		+
SEQ ID No. 1226	2849.1	Contig45	p	39757	40530	-	-	93%	+
SEQ ID No. 711	2029.2	Contig45	m	41434	42672		-	95%	+
SEQ ID No. 710	2027.2	Contig45	m	42859	44145	-	+		-
SEQ ID No. 709	2026.1	Contig45	p	44695	44994	-	-	80%	+
SEQ ID No. 2102	417.3	Contig45	m	45095	45925	-	+		+
SEQ ID No. 2087	415.2	Contig45	p	46532	47668	-	+		-
SEQ ID No. 2069	413.5	Contig45	m	48015	50384	-	+		-
SEQ ID No. 2623	4927.1	Contig45	m	50814	51575	-	+		+
SEQ ID No. 2624	4929.2	Contig45	m	51754	52683	-	-	98%	-
SEQ ID No. 2626	4930.2	Contig45	m	52692	52979	-	-	100%	-
SEQ ID No. 2404	4604.3	Contig45	m	53037	53615	-	-	99%	-
SEQ ID No. 3306	790.3	Contig45	m	53670	56318	-	-	95%	+
SEQ ID No. 1259	2915.2	Contig45	m	56619	59432	-	-	97%	+
SEQ ID No. 1915	390.2	Contig45	p	59678	60028	-	-	99%	-
SEQ ID No. 1907	389.3	Contig45	p	59965	61974	-	-	98%	-
SEQ ID No. 1258	2914.1	Contig45	p	61952	62623	-	-	97%	-
SEQ ID No. 482	1655.2	Contig45	p	62532	62795	-	-	100%	-
SEQ ID No. 483	1657.1	Contig45	p	62789	63544	-	-	98%	-
SEQ ID No. 3035	5730.1	Contig45	m	63648	63923		-	97%	+
SEQ ID No. 484	1658.2	Contig45	p	63762	65222	-	-	98%	+
SEQ ID No. 153	1142.4	Contig45	m	65328	66779	+	-	99%	-

SEQ ID No. 1257	2911.3	Contig45	m	66824	67198	-	-	99%	-
SEQ ID No. 795	2169.2	Contig45	p	67426	67854	+	-	97%	-
SEQ ID No. 796	2170.1	Contig45	p	68018	68461	+	-	97%	-
SEQ ID No. 3449	993.2	Contig45	m	68458	69726	-	-	99%	-
SEQ ID No. 3448	992.2	Contig45	p	68996	69385		-	100%	+
SEQ ID No. 3450	994.2	Contig45	m	69800	72601	+	-	99%	-
SEQ ID No. 1256	2910.1	Contig45	p	72535	74043	-	-	98%	-
SEQ ID No. 1254	2909.1	Contig45	p	74272	75552	-	-	98%	-
SEQ ID No. 1253	2908.2	Contig45	p	75581	76588	-	-	98%	+
SEQ ID No. 1252	2906.2	Contig45	p	76746	77390	-	-	97%	-
SEQ ID No. 1251	2905.2	Contig45	p	77708	79522		-	98%	+
SEQ ID No. 241	1279.3	Contig45	p	79527	80066	-	-	96%	-
SEQ ID No. 240	1278.2	Contig45	m	80071	81282		-	99%	-
SEQ ID No. 754	2108.2	Contig45	p	81438	82382		-	99%	-
SEQ ID No. 102	1069.2	Contig45	p	82510	83586	-	+		-
SEQ ID No. 101	1067.1	Contig45	p	83801	84055		+		-
SEQ ID No. 100	1066.2	Contig45	m	84221	85126	-	-	94%	-
SEQ ID No. 1735	3608.1	Contig45	m	85559	85888	-	-	100%	+
SEQ ID No. 1734	3607.2	Contig45	m	85993	87189	-	-	97%	-
SEQ ID No. 2763	5173.1	Contig45	m	87262	89307		-	96%	-
SEQ ID No. 1842	3765.1	Contig45	m	89250	90209	-	-	98%	-
SEQ ID No. 958	2429.2	Contig45	m	90197	91141	-	-	99%	-
SEQ ID No. 957	2428.2	Contig45	p	91233	91619	-	+		-
SEQ ID No. 956	2427.4	Contig45	m	91643	92011	-	+		-
SEQ ID No. 1841	3764.3	Contig45	p	92316	92675	-	+		+
SEQ ID No. 1840	3763.1	Contig45	p	92636	93379	-	+		-
SEQ ID No. 1839	3761.1	Contig45	p	93268	93459		+		+
SEQ ID No. 1838	3760.3	Contig45	m	93555	95363	+	-	99%	-
SEQ ID No. 2032	4075.3	Contig45	m	95380	95859	-	-	98%	-
SEQ ID No. 2033	4076.1	Contig45	m	96123	96674	-	-	95%	+
SEQ ID No. 534	1737.2	Contig45	p	96822	97934	-	+		-
SEQ ID No. 533	1735.1	Contig45	p	98103	98318	-	+		+
SEQ ID No. 532	1734.2	Contig45	p	98319	98483		+		+
SEQ ID No. 2034	4078.1	Contig45	m	98517	99893	-	-	97%	-
SEQ ID No. 2036	4080.2	Contig45	p	100247	101554	+	-	97%	-
SEQ ID No. 2764	5174.2	Contig45	m	101564	102016	+	-	99%	+

SEQ ID No. 496	1675.3	Contig45	m	102037	102357		-	99%	-
SEQ ID No. 495	1673.5	Contig45	p	102412	104136	-	-	97%	-
SEQ ID No. 492	1669.3	Contig45	m	104200	105201	-	-	98%	-
SEQ ID No. 493	1670.3	Contig45	p	105391	106725	-	-	96%	-
SEQ ID No. 2135	4222.3	Contig45	m	106907	107263	-	-	97%	-
SEQ ID No. 2134	4221.3	Contig45	p	107390	107692	-	-	97%	-
SEQ ID No. 2133	4220.2	Contig45	m	107870	108175	-	+		-
SEQ ID No. 1884	384.3	Contig45	m	108189	109274	-	+		-
SEQ ID No. 502	1685.3	Contig45	m	109460	110584	-	-	98%	-
SEQ ID No. 3390	914.3	Contig45	m	110599	112401	-	-	98%	-
SEQ ID No. 3391	915.3	Contig45	m	112344	113216	+	-	97%	-
SEQ ID No. 2765	5176.1	Contig45	m	113642	113920		+		+
SEQ ID No. 2766	5177.2	Contig45	p	114007	114765	-	-	99%	-
SEQ ID No. 3036	5735.2	Contig45	p	114761	115966	+	-	98%	-
SEQ ID No. 2972	5598.2	Contig45	p	116104	117363	+	-	99%	-
SEQ ID No. 3166	6098.1	Contig45	p	117531	117845	-	+		-
SEQ ID No. 2256	4387.3	Contig45	p	117746	118360		-	93%	-
SEQ ID No. 2257	4388.1	Contig45	m	118391	119677	-	-	100%	-
SEQ ID No. 87	1044.3	Contig45	m	119649	120962	-	-	99%	-
SEQ ID No. 86	1042.5	Contig45	m	121153	122400	-	-	100%	-
SEQ ID No. 2973	5600.2	Contig45	m	122515	122787	-	-	100%	-
SEQ ID No. 3082	588.3	Contig45	m	122879	123670	-	-	96%	-
SEQ ID No. 3076	587.2	Contig45	m	123642	124427	-	-	98%	-
SEQ ID No. 3072	585.2	Contig45	m	124480	124830	-	-	99%	-
SEQ ID No. 3068	583.2	Contig45	m	125324	125962	-	-	98%	-
SEQ ID No. 3165	6097.1	Contig45	m	126102	134465		+		-
SEQ ID No. 3164	6094.1	Contig46	m	94	366		+		+
SEQ ID No. 2413	4620.1	Contig46	p	136	441	-	+		-
SEQ ID No. 2411	4618.1	Contig46	p	588	1376	-	-	100%	-
SEQ ID No. 2410	4616.2	Contig46	p	1343	2314	+	-	99%	-
SEQ ID No. 3012	5672.1	Contig46	p	2296	2640	+	+		+
SEQ ID No. 2409	4615.3	Contig46	m	2641	3060		-	100%	+
SEQ ID No. 2907	5479.2	Contig46	m	3146	3784	-	-	100%	-
SEQ ID No. 1830	3747.2	Contig46	m	3742	4659	+	-	100%	-
SEQ ID No. 1831	3748.1	Contig46	m	4653	6005	-	-	100%	-
SEQ ID No. 1832	3750.2	Contig46	p	6131	7858	+	-	100%	-

SEQ ID No. 1833	3752.2	Contig46	p	7797	8543	+	-	100%		-
SEQ ID No. 1834	3753.2	Contig46	p	8556	10478	+	-	100%		-
SEQ ID No. 1835	3754.2	Contig46	p	10557	11339	+	-	100%		-
SEQ ID No. 1836	3756.2	Contig46	p	11314	12657	-	-	100%		-
SEQ ID No. 2306	4469.2	Contig46	m	12750	13640	-	-	100%		-
SEQ ID No. 3308	793.2	Contig46	m	13696	14010		-	100%		+
SEQ ID No. 3309	794.2	Contig46	m	13863	14069		+			+
SEQ ID No. 3310	795.2	Contig46	m	14074	14865	-	-	100%		-
SEQ ID No. 2308	4470.1	Contig46	m	14843	15916	-	-	100%		-
SEQ ID No. 352	1457.2	Contig46	m	15855	16466	-	-	98%		-
SEQ ID No. 351	1456.2	Contig46	m	16467	17201	-	-	100%		-
SEQ ID No. 350	1455.2	Contig46	m	17189	17725	+	-	100%		-
SEQ ID No. 2309	4471.3	Contig46	m	17679	18281	+	-	100%		-
SEQ ID No. 3116	5978.2	Contig46	m	18245	18802	-	-	100%		-
SEQ ID No. 3095	5903.2	Contig46	m	18786	19751	-	-	100%	+	-
SEQ ID No. 3024	5696.2	Contig46	p	20113	20913	-	-	100%	+	-
SEQ ID No. 3023	5695.1	Contig46	p	20906	21688	-	-	100%		-
SEQ ID No. 3022	5694.2	Contig46	p	21689	22165	+	-	100%		-
SEQ ID No. 2912	5489.2	Contig46	p	22170	22781	+	-	100%		-
SEQ ID No. 2219	4341.3	Contig46	p	22769	23065	-	-	100%		+
SEQ ID No. 2220	4342.1	Contig46	p	23446	23700	-	+			-
SEQ ID No. 2221	4344.1	Contig46	p	23660	24961	-	-	100%		-
SEQ ID No. 2222	4345.1	Contig46	p	24955	25716	-	-	100%		-
SEQ ID No. 366	1484.2	Contig46	m	25905	26603	-	-	100%		-
SEQ ID No. 367	1485.4	Contig46	m	26579	27880	+	-	100%		-
SEQ ID No. 2573	4848.3	Contig46	m	27829	28851	+	-	100%		-
SEQ ID No. 2572	4847.2	Contig46	p	29223	29828		-	100%		+
SEQ ID No. 123	1096.3	Contig46	p	30003	31439	-	-	100%		-
SEQ ID No. 122	1094.1	Contig46	m	31483	31767	-	-	100%		-
SEQ ID No. 121	1093.3	Contig46	m	31987	32967	-	-	100%		-
SEQ ID No. 2376	4564.2	Contig46	p	32890	33672	-	-	100%		-
SEQ ID No. 1069	2597.4	Contig46	p	33657	34349	-	-	100%		-
SEQ ID No. 1433	3159.2	Contig46	p	34524	35318	+	-	100%		-
SEQ ID No. 1432	3158.1	Contig46	p	35182	35460	-	+			-
SEQ ID No. 1431	3157.1	Contig46	p	35466	35897	+	-	100%		-
SEQ ID No. 1430	3153.3	Contig46	p	35885	37846	+	-	100%		-

SEQ ID No. 1429	3152.1	Contig46	p	37789	38376	+	-	100%		-
SEQ ID No. 744	2085.2	Contig46	p	38340	38777	+	-	100%		-
SEQ ID No. 743	2083.2	Contig46	p	38725	39456	+	-	100%		-
SEQ ID No. 742	2082.2	Contig46	p	39407	39649	-	-	100%		-
SEQ ID No. 614	1853.2	Contig46	p	39690	40109	+	-	100%		-
SEQ ID No. 3454	998.3	Contig46	m	40865	42739	-	-	100%	+	-
SEQ ID No. 3455	999.2	Contig46	p	42796	43959	-	-	100%	+	-
SEQ ID No. 58	1000.2	Contig46	p	43955	44752	-	-	100%	+	-
SEQ ID No. 406	1542.2	Contig46	p	44743	45813	-	-	100%		-
SEQ ID No. 404	1539.2	Contig46	m	45979	46878	+	-	100%		+
SEQ ID No. 209	1230.2	Contig46	p	47057	48649	-	-	100%		-
SEQ ID No. 210	1231.3	Contig46	m	48715	49140	-	-	100%		+
SEQ ID No. 2522	4773.3	Contig46	m	49233	50660	+	-	100%		-
SEQ ID No. 2523	4774.1	Contig46	m	50654	50962	-	-	100%		-
SEQ ID No. 2524	4775.2	Contig46	m	50943	52103	-	-	100%		-
SEQ ID No. 2525	4776.3	Contig46	m	52470	53033	-	-	100%		-
SEQ ID No. 3163	6093.1	Contig46	m	53278	53583	-	-	100%		+
SEQ ID No. 728	2060.3	Contig46	p	53835	54983	-	-	100%		-
SEQ ID No. 729	2062.2	Contig46	m	55045	55689	+	-	100%		+
SEQ ID No. 731	2065.2	Contig46	p	56118	56603	-	-	100%		+
SEQ ID No. 730	2064.3	Contig46	m	56324	56575	-	-	100%		+
SEQ ID No. 732	2066.5	Contig46	p	56724	57173	-	-	100%		+
SEQ ID No. 733	2067.5	Contig46	p	57300	58376	-	-	100%		-
SEQ ID No. 2026	4068.2	Contig46	p	58873	60171	-	-	100%		-
SEQ ID No. 2025	4067.2	Contig46	m	59556	59759		-	100%		+
SEQ ID No. 2190	4303.1	Contig46	m	60495	61640	-	-	99%		-
SEQ ID No. 2189	4302.1	Contig46	m	61621	62454	-	-	100%		-
SEQ ID No. 2188	4301.1	Contig46	m	62455	63360	-	-	100%		-
SEQ ID No. 332	1420.2	Contig46	m	63415	64569	-	-	100%		-
SEQ ID No. 333	1421.3	Contig46	p	64694	67030	-	-	100%		-
SEQ ID No. 2888	5446.3	Contig46	p	67332	68681	-	-	100%		-
SEQ ID No. 2889	5447.1	Contig46	p	68866	69651	-	-	100%		-
SEQ ID No. 600	1834.4	Contig46	p	69560	70978	-	-	100%		-
SEQ ID No. 3117	5979.1	Contig46	m	70998	71240		-	100%		+
SEQ ID No. 2890	5448.1	Contig46	p	71292	71837	-	-	100%		-
SEQ ID No. 2380	4571.2	Contig46	p	71819	72220	-	-	100%		-

SEQ ID No. 1055	2576.2	Contig46	m	72302	73279	+	-	100%		-
SEQ ID No. 1056	2578.2	Contig46	m	73534	74286	-	-	100%		+
SEQ ID No. 2381	4572.2	Contig46	m	74477	75556	-	-	100%		-
SEQ ID No. 2382	4573.4	Contig46	m	75541	77133	+	-	100%		+
SEQ ID No. 3096	5904.2	Contig46	p	77481	78152	-	-	100%		-
SEQ ID No. 2673	5010.2	Contig46	p	78136	79059	+	-	100%		-
SEQ ID No. 2674	5015.2	Contig46	p	79272	79613	-	-	100%		-
SEQ ID No. 2896	5460.1	Contig46	p	79588	80679	+	-	99%		-
SEQ ID No. 1959	3974.2	Contig46	p	80784	81344	-	-	100%		-
SEQ ID No. 1958	3972.2	Contig46	p	81539	82297	-	-	100%	+	-
SEQ ID No. 1957	3971.2	Contig46	m	82354	82878	-	-	100%		-
SEQ ID No. 1956	3970.1	Contig46	m	82868	83266	-	-	100%		-
SEQ ID No. 1954	3969.1	Contig46	m	83279	83998	+	-	100%		-
SEQ ID No. 1953	3968.1	Contig46	p	84058	84474	-	-	100%		-
SEQ ID No. 1952	3967.1	Contig46	p	84548	85108	+	-	100%		-
SEQ ID No. 1951	3966.1	Contig46	p	85386	86165	-	-	100%		-
SEQ ID No. 1950	3965.1	Contig46	p	86726	87241	-	-	100%		-
SEQ ID No. 1949	3964.1	Contig46	p	87172	88179	-	-	100%		-
SEQ ID No. 1948	3962.2	Contig46	p	88158	88514	-	-	100%		-
SEQ ID No. 2897	5462.2	Contig46	p	88889	90562	-	-	99%		-
SEQ ID No. 3205	643.3	Contig46	p	90559	92034	-	-	100%		-
SEQ ID No. 3206	644.2	Contig46	m	92079	93134	-	-	100%		-
SEQ ID No. 3207	645.4	Contig46	m	93128	93838	-	-	100%		-
SEQ ID No. 3208	647.5	Contig46	p	94033	94683		-	100%		-
SEQ ID No. 2579	4859.2	Contig46	p	94693	95967	-	-	99%		+
SEQ ID No. 2581	4860.1	Contig46	p	95971	96123		+			+
SEQ ID No. 2924	5515.2	Contig46	p	96370	97125	-	-	99%		-
SEQ ID No. 1939	3942.2	Contig46	p	97130	98077	-	-	99%		-
SEQ ID No. 3191	623.2	Contig46	p	98038	99210	-	-	98%		-
SEQ ID No. 3190	622.3	Contig46	m	99282	101681	+	-	99%		-
SEQ ID No. 2923	5514.2	Contig46	m	101816	102838		-	89%		+
SEQ ID No. 2865	5398.3	Contig46	p	103240	104304	-	-	99%		-
SEQ ID No. 1697	3555.2	Contig46	p	104305	104856	-	-	99%		-
SEQ ID No. 1696	3554.4	Contig46	p	104831	105469	-	-	99%		-
SEQ ID No. 642	1895.6	Contig46	p	105421	106050		-	96%		-
SEQ ID No. 643	1896.5	Contig46	p	106051	108150	+	-	99%		-

SEQ ID No. 1695	3551.1	Contig46	p	108590	109123	+	-	100%		-
SEQ ID No. 1694	3550.1	Contig46	p	109101	110219	-	-	98%		-
SEQ ID No. 1693	3548.2	Contig46	p	110195	111667	-	-	97%		+
SEQ ID No. 1692	3547.1	Contig46	p	111746	112207	-	-	99%		-
SEQ ID No. 1691	3545.1	Contig46	p	112319	113317	-	-	98%		-
SEQ ID No. 370	1490.4	Contig46	p	113424	116225	-	-	99%		-
SEQ ID No. 369	1488.2	Contig46	m	114650	114994		-	98%		+
SEQ ID No. 830	2233.3	Contig46	p	116207	116686	+	-	98%		-
SEQ ID No. 829	2232.1	Contig46	m	116985	117551	-	-	98%		-
SEQ ID No. 242	1280.3	Contig46	m	117725	119929	-	-	98%		+
SEQ ID No. 292	1366.3	Contig46	m	120156	123287	-	-	97%		+
SEQ ID No. 1750	3629.2	Contig46	p	123454	124431	-	-	99%		-
SEQ ID No. 1455	3193.2	Contig46	p	124409	125110	-	-	98%		-
SEQ ID No. 1454	3191.1	Contig46	m	125166	126344	-	-	97%		+
SEQ ID No. 1453	3190.1	Contig46	m	126503	127390	-	-	94%		-
SEQ ID No. 788	2159.2	Contig46	p	127607	128428	-	-	99%		-
SEQ ID No. 455	1615.2	Contig46	p	128397	130232	-	-	99%	+	-
SEQ ID No. 790	2160.2	Contig46	p	130192	131184	-	-	98%	+	-
SEQ ID No. 2111	418.4	Contig46	m	131572	132249	+	-	100%		-
SEQ ID No. 2116	419.2	Contig46	p	132432	133271	-	-	98%		-
SEQ ID No. 2128	421.2	Contig46	p	133283	134725	-	-	99%	+	-
SEQ ID No. 2132	422.3	Contig46	p	134726	135649	-	-	98%		-
SEQ ID No. 605	1841.2	Contig46	p	135667	137343	-	-	98%		-
SEQ ID No. 1452	3186.2	Contig46	m	137322	140789	-	-	99%	+	-
SEQ ID No. 1451	3184.3	Contig46	m	140860	142122	-	-	99%		-
SEQ ID No. 89	1048.3	Contig46	m	142126	143316		-	99%		-
SEQ ID No. 88	1047.2	Contig46	m	143301	144233	+	-	98%		-
SEQ ID No. 886	2330.2	Contig46	p	144380	146443	-	-	98%		-
SEQ ID No. 887	2333.3	Contig46	p	146545	148728	-	-	98%		-
SEQ ID No. 2683	5030.2	Contig46	p	148587	149702	-	-	99%		-
SEQ ID No. 1931	3927.2	Contig46	p	149681	151072	-	-	98%		-
SEQ ID No. 469	1636.2	Contig46	p	151141	154593	-	-	99%		-
SEQ ID No. 1930	3925.2	Contig46	m	154679	155944	-	-	99%		+
SEQ ID No. 1929	3924.2	Contig46	p	156244	157134	-	-	98%		-
SEQ ID No. 243	1282.3	Contig46	p	157074	158642	-	-	97%		-
SEQ ID No. 1928	3923.2	Contig46	p	158658	160574	-	-	99%		-

SEQ ID No. 2570	4844.2	Contig46	m	160720	161328	-	-	93%		+
SEQ ID No. 2571	4845.4	Contig46	m	161597	163333	-	-	96%		+
SEQ ID No. 2161	4255.1	Contig46	m	163658	164020		-	100%		+
SEQ ID No. 2160	4254.2	Contig46	p	163752	164525	-	-	96%		+
SEQ ID No. 2159	4253.1	Contig46	m	164822	166306	-	-	100%		-
SEQ ID No. 873	2309.2	Contig46	p	166657	167811	+	-	97%		-
SEQ ID No. 872	2308.2	Contig46	p	168348	168671		+			+
SEQ ID No. 3162	6089.1	Contig46	m	168652	169203	-	+			-
SEQ ID No. 3120	5985.1	Contig47	m	21	398		+			-
SEQ ID No. 3083	5890.1	Contig47	m	596	778	-	+			+
SEQ ID No. 1768	3653.3	Contig47	m	801	1010	-	+			-
SEQ ID No. 1769	3655.1	Contig47	m	1280	1993	-	+			+
SEQ ID No. 2733	510.2	Contig47	m	1971	2675		+			+
SEQ ID No. 2726	509.1	Contig47	m	2774	3367		+			-
SEQ ID No. 2713	507.3	Contig47	m	3397	3723	-	-	93%		-
SEQ ID No. 2704	506.3	Contig47	p	4108	4518	-	+			+
SEQ ID No. 2697	505.3	Contig47	m	4559	5503	-	+			+
SEQ ID No. 1770	3657.1	Contig47	m	5811	6452	-	+			-
SEQ ID No. 1771	3658.1	Contig47	p	6810	7100		+			+
SEQ ID No. 3417	948.3	Contig47	p	7066	7596	-	+			-
SEQ ID No. 3416	947.3	Contig47	m	7565	7927		+			-
SEQ ID No. 3415	946.3	Contig47	m	7965	8639	-	+			-
SEQ ID No. 1772	3659.2	Contig47	m	8757	9299	-	+			+
SEQ ID No. 1774	3661.4	Contig47	m	9263	10255	+	+			-
SEQ ID No. 2984	5620.3	Contig47	p	10879	11130	-	-	96%		-
SEQ ID No. 786	2153.3	Contig47	p	11124	11363		+			+
SEQ ID No. 2163	426.3	Contig47	p	11669	12370	-	+			+
SEQ ID No. 2156	425.1	Contig47	m	12661	12954	-	+			+
SEQ ID No. 2148	424.1	Contig47	m	12935	13252	-	-	96%		-
SEQ ID No. 2140	423.2	Contig47	m	13262	14566	+	+			-
SEQ ID No. 2985	5621.1	Contig47	m	14589	15851		+			+
SEQ ID No. 2986	5623.2	Contig47	m	16093	17499	-	+			+
SEQ ID No. 1909	3891.2	Contig47	m	17492	18145	-	+		+	-
SEQ ID No. 1908	3890.2	Contig47	p	18284	19642	-	+			+
SEQ ID No. 1906	3889.1	Contig47	m	19632	20417		+			-
SEQ ID No. 1905	3888.1	Contig47	m	20390	20593		+			+

SEQ ID No. 1904	3887.1	Contig47	m	20607	21833		+			+
SEQ ID No. 1903	3884.2	Contig47	p	22460	22942	-	-	96%		+
SEQ ID No. 338	1432.3	Contig47	p	23126	24049	-	-	98%		-
SEQ ID No. 337	1429.4	Contig47	m	24232	25806	+	+			+
SEQ ID No. 2868	5404.2	Contig47	m	26133	27047	-	+			+
SEQ ID No. 2867	5402.1	Contig47	p	27185	28633	-	-	86%		+
SEQ ID No. 1028	2533.3	Contig47	p	28867	29787	-	-	84%		-
SEQ ID No. 1027	2530.3	Contig47	p	29880	31814	-	-	95%		-
SEQ ID No. 1733	3602.2	Contig47	p	31894	33795	-	-	87%		-
SEQ ID No. 1732	3601.1	Contig47	p	33973	35130	-	+			-
SEQ ID No. 1731	3600.2	Contig47	p	35321	36292	-	+			-
SEQ ID No. 975	2453.4	Contig47	p	36235	39030	-	+			-
SEQ ID No. 480	1652.4	Contig47	p	39121	39687	-	-	91%		-
SEQ ID No. 481	1653.3	Contig47	m	39802	40740	+	-	96%		-
SEQ ID No. 3148	606.3	Contig47	m	40856	42601	+	-	95%	+	-
SEQ ID No. 3174	612.2	Contig47	p	42746	44959	-	-	96%		-
SEQ ID No. 740	2079.2	Contig47	m	45270	46547	-	-	93%		+
SEQ ID No. 739	2078.2	Contig47	m	46681	46965		+			+
SEQ ID No. 1783	3673.2	Contig47	m	47107	48360	-	-	94%		-
SEQ ID No. 3057	5809.1	Contig47	p	48148	48435		-	79%		+
SEQ ID No. 2617	4917.2	Contig47	m	48380	49735	-	-	94%		+
SEQ ID No. 2616	4916.2	Contig47	p	49917	50837		-	95%		-
SEQ ID No. 3004	5655.2	Contig47	m	50958	51446	-	-	99%		-
SEQ ID No. 3003	5654.2	Contig47	p	51771	52181	-	-	94%		+
SEQ ID No. 2628	4935.2	Contig47	m	52309	53451	-	-	97%		-
SEQ ID No. 2627	4934.3	Contig47	m	53749	54948	-	-	98%		-
SEQ ID No. 3056	5808.1	Contig47	m	55003	55581	-	-	97%		-
SEQ ID No. 3427	964.3	Contig47	m	55765	57123	-	-	90%		+
SEQ ID No. 3421	953.2	Contig47	p	57651	59225	-	-	89%		-
SEQ ID No. 3420	952.1	Contig47	m	59202	60149	-	-	98%		-
SEQ ID No. 3419	950.2	Contig47	m	60306	61823	-	-	85%		-
SEQ ID No. 1493	3252.1	Contig47	p	62076	63032	-	-	98%		+
SEQ ID No. 1492	3251.1	Contig47	p	62996	63430	-	-	99%		-
SEQ ID No. 3383	901.2	Contig47	m	63514	64692	-	-	87%		+
SEQ ID No. 3384	903.2	Contig47	p	64977	66098	-	-	94%		+
SEQ ID No. 3385	905.2	Contig47	m	66185	67045	+	+			-

SEQ ID No. 1491	3250.1	Contig47	m	67309	67623	-	-	96%	+
SEQ ID No. 1489	3249.2	Contig47	p	67801	68496	-	-	98%	-
SEQ ID No. 418	1562.3	Contig47	p	68820	69641	-	-	97%	+
SEQ ID No. 3161	6087.1	Contig47	m	69577	69876		+		+
SEQ ID No. 417	1560.2	Contig47	m	69761	70600		+		-
SEQ ID No. 1488	3248.1	Contig47	m	70560	71483	-	+		-
SEQ ID No. 1487	3246.2	Contig47	p	71805	73802	-	-	95%	-
SEQ ID No. 619	1861.2	Contig47	m	73905	74735	-	-	96%	+
SEQ ID No. 618	1860.2	Contig47	p	75237	76346		-	92%	-
SEQ ID No. 1546	3336.2	Contig47	p	76579	77034	+	-	93%	+
SEQ ID No. 1547	3337.2	Contig47	m	77247	77543	-	-	90%	+
SEQ ID No. 1548	3338.1	Contig47	p	77766	78158	-	+		+
SEQ ID No. 1550	3341.2	Contig47	p	78359	81394	-	+		+
SEQ ID No. 1551	3343.1	Contig47	m	81608	82249		+		+
SEQ ID No. 1552	3345.2	Contig47	m	82458	84782		-	98%	+
SEQ ID No. 1062	2585.2	Contig47	p	84972	86738	+	-	95%	+
SEQ ID No. 422	1567.2	Contig47	m	86827	88173	-	-	92%	+
SEQ ID No. 421	1566.3	Contig47	m	88505	89068	+	+		-
SEQ ID No. 490	1664.2	Contig47	m	89346	90161	-	-	87%	+
SEQ ID No. 2818	5297.2	Contig47	p	90488	91873	-	+		+
SEQ ID No. 2345	4519.3	Contig47	p	92237	93889	-	-	97%	-
SEQ ID No. 3438	979.3	Contig47	p	93871	95094	-	-	94%	-
SEQ ID No. 3439	980.1	Contig47	p	95095	95814	-	-	98%	-
SEQ ID No. 3440	981.2	Contig47	p	95787	96434		-	84%	-
SEQ ID No. 2344	4517.2	Contig47	p	96391	97248	-	-	92%	-
SEQ ID No. 2558	4819.2	Contig47	p	97143	98192	-	-	93%	+
SEQ ID No. 2560	4821.2	Contig47	p	98266	99219	-	-	94%	-
SEQ ID No. 2575	4850.2	Contig47	p	99385	101181	-	-	93%	+
SEQ ID No. 2576	4851.2	Contig47	p	101415	101942	-	-	94%	-
SEQ ID No. 897	2347.2	Contig47	p	101957	103012		-	92%	+
SEQ ID No. 896	2346.1	Contig47	p	103046	103429		-	91%	+
SEQ ID No. 895	2345.2	Contig47	m	103536	103991		-	96%	+
SEQ ID No. 275	1336.3	Contig47	m	104195	105385	-	-	98%	-
SEQ ID No. 2136	4225.1	Contig47	p	105500	105769		-	94%	+
SEQ ID No. 881	2321.3	Contig47	m	105652	106263	-	-	98%	-
SEQ ID No. 2871	5412.1	Contig47	p	106268	107311	-	-	97%	-

SEQ ID No. 2395	4591.2	Contig47	p	107343	107822	-	-	94%	-
SEQ ID No. 2394	4590.1	Contig47	p	107862	108332	-	-	100%	-
SEQ ID No. 2392	4589.1	Contig47	m	108456	109088	-	-	96%	+
SEQ ID No. 2391	4588.2	Contig47	m	109240	109881	-	-	99%	+
SEQ ID No. 2390	4587.2	Contig47	p	109974	110486		-	97%	+
SEQ ID No. 2872	5413.1	Contig47	p	110505	110849	-	-	100%	+
SEQ ID No. 949	2419.3	Contig47	p	110960	111712	-	-	98%	-
SEQ ID No. 948	2418.2	Contig47	p	111783	112523	+	-	97%	-
SEQ ID No. 947	2415.2	Contig47	m	112603	113577	-	-	95%	-
SEQ ID No. 2103	4170.1	Contig47	m	113730	116510	-	-	97%	-
SEQ ID No. 2101	4168.2	Contig47	p	116698	117213	-	-	98%	+
SEQ ID No. 2529	4780.2	Contig47	m	117261	117761		-	95%	+
SEQ ID No. 2528	4779.1	Contig47	p	118015	118479	-	-	97%	-
SEQ ID No. 2527	4778.1	Contig47	p	118604	119269	-	-	88%	+
SEQ ID No. 2526	4777.2	Contig47	p	119435	120775	-	-	96%	+
SEQ ID No. 2873	5417.1	Contig47	m	120900	121331	-	-	99%	-
SEQ ID No. 2233	4357.3	Contig47	m	121371	122663	-	-	97%	-
SEQ ID No. 2234	4358.2	Contig47	m	122641	123207	-	-	96%	-
SEQ ID No. 2964	5586.1	Contig47	m	123420	123806	-	-	96%	-
SEQ ID No. 998	2485.3	Contig47	p	123935	124582	-	-	97%	+
SEQ ID No. 629	1878.3	Contig47	p	124766	125719	-	-	98%	-
SEQ ID No. 628	1875.2	Contig47	p	125840	127069	+	+		-
SEQ ID No. 2236	4363.2	Contig47	p	127205	127927		-	98%	-
SEQ ID No. 2717	5076.2	Contig47	m	127987	128469	-	-	87%	-
SEQ ID No. 2718	5077.3	Contig47	m	128463	129851	-	-	99%	-
SEQ ID No. 1767	3652.2	Contig47	m	129748	130578	-	-	98%	-
SEQ ID No. 1766	3651.1	Contig47	m	130551	131399	-	-	98%	-
SEQ ID No. 1765	3650.1	Contig47	m	131389	132657		-	97%	-
SEQ ID No. 1011	2506.2	Contig47	m	132564	133628	-	-	98%	-
SEQ ID No. 1012	2508.3	Contig47	m	133619	134782	-	-	98%	-
SEQ ID No. 1763	3647.2	Contig47	m	134722	134958	-	+		-
SEQ ID No. 264	1320.3	Contig47	m	134959	137187	-	-	97%	-
SEQ ID No. 1762	3645.3	Contig47	m	137198	137989	-	-	94%	-
SEQ ID No. 3055	5807.1	Contig47	m	137956	138348	-	-	99%	-
SEQ ID No. 2471	4704.2	Contig47	m	138309	139100	-	-	98%	-
SEQ ID No. 2470	4703.2	Contig47	p	139054	139539		+		-

SEQ ID No. 2469	4702.1	Contig47	p	139494	140090		-	94%	-
SEQ ID No. 2468	4701.2	Contig47	m	140087	140926	-	-	93%	-
SEQ ID No. 2507	4752.2	Contig47	m	141084	142007	-	-	98%	-
SEQ ID No. 2508	4754.1	Contig47	p	142255	143601	-	-	97%	+
SEQ ID No. 2509	4755.2	Contig47	m	143628	144191	-	-	98%	-
SEQ ID No. 2024	4066.2	Contig47	p	144222	144818	-	-	96%	-
SEQ ID No. 2023	4065.3	Contig47	p	144790	146553	-	-	98%	-
SEQ ID No. 2022	4063.1	Contig47	m	146659	148047	-	-	97%	-
SEQ ID No. 2021	4061.1	Contig47	m	148048	148587	-	-	99%	-
SEQ ID No. 2020	4060.1	Contig47	m	148533	149945	-	-	97%	-
SEQ ID No. 488	1661.3	Contig47	p	149794	150489	-	-	98%	-
SEQ ID No. 487	1660.2	Contig47	p	150510	150761		-	89%	+
SEQ ID No. 485	1659.4	Contig47	p	150829	154725	-	-	96%	+
SEQ ID No. 835	2242.1	Contig47	m	154849	155424	-	-	98%	+
SEQ ID No. 312	1394.2	Contig47	m	155492	155962	-	+		-
SEQ ID No. 313	1396.3	Contig47	p	156048	157058	-	-	98%	-
SEQ ID No. 314	1397.2	Contig47	p	157308	157814	-	-	99%	-
SEQ ID No. 1867	3811.1	Contig47	m	157863	158390	-	+		-
SEQ ID No. 1866	3810.1	Contig47	p	158615	159328	-	-	97%	-
SEQ ID No. 587	1807.3	Contig47	m	159379	160806	-	-	100%	-
SEQ ID No. 586	1806.2	Contig47	p	160979	161524	-	-	99%	-
SEQ ID No. 585	1805.3	Contig47	m	161682	162344	-	-	98%	-
SEQ ID No. 2368	4551.2	Contig47	m	162316	162594	-	-	100%	-
SEQ ID No. 2369	4552.4	Contig47	p	162799	163758	-	-	89%	+
SEQ ID No. 2844	5354.3	Contig47	p	164077	164319	+	-	98%	+
SEQ ID No. 356	1467.5	Contig47	p	164422	165021	-	-	98%	-
SEQ ID No. 355	1466.2	Contig47	p	165058	165801	-	-	98%	-
SEQ ID No. 2430	4645.2	Contig47	p	165886	167118	-	-	99%	-
SEQ ID No. 2429	4644.3	Contig47	m	167115	168464	-	-	95%	+
SEQ ID No. 2126	4206.2	Contig47	p	168685	171516	-	-	97%	+
SEQ ID No. 2125	4205.2	Contig47	p	171437	172408	-	-	91%	+
SEQ ID No. 2124	4203.2	Contig47	p	172808	174076	-	-	100%	-
SEQ ID No. 2123	4200.2	Contig47	m	174124	174822		-	99%	+
SEQ ID No. 301	1377.5	Contig47	m	175216	177651	-	-	97%	+
SEQ ID No. 302	1378.5	Contig47	m	177711	178910	-	-	98%	+
SEQ ID No. 2843	5349.3	Contig47	p	179148	181841	-	+		+

SEQ ID No. 2286	4431.2	Contig47	p	181989	184718	-	-	91%		+
SEQ ID No. 2746	5129.4	Contig47	m	185037	188228	+	-	99%		-
SEQ ID No. 2589	4873.3	Contig47	m	188249	189448	-	-	98%		-
SEQ ID No. 2588	4872.3	Contig47	m	189409	190986	+	-	98%		-
SEQ ID No. 3160	6086.1	Contig47	m	191153	191521	-	-	98%		-
SEQ ID No. 3016	5685.2	Contig47	m	191522	192799	-	-	98%		-
SEQ ID No. 3017	5686.2	Contig47	p	192951	193427	-	-	98%		-
SEQ ID No. 3018	5687.1	Contig47	p	193621	193965	-	-	99%		+
SEQ ID No. 3019	5688.2	Contig47	p	194152	194715		-	79%		+
SEQ ID No. 2958	5569.1	Contig47	m	194808	195188	+	-	98%		+
SEQ ID No. 2957	5566.1	Contig47	p	195344	195577	+	+			-
SEQ ID No. 99	1065.3	Contig47	m	195618	196814	-	-	96%		-
SEQ ID No. 98	1063.1	Contig47	m	197173	198501	-	-	98%		+
SEQ ID No. 2664	4998.2	Contig47	m	198880	200607	-	-	96%		+
SEQ ID No. 2663	4995.1	Contig47	p	200711	201430	-	-	98%		+
SEQ ID No. 1604	3418.1	Contig48	p	1	852		-	96%		+
SEQ ID No. 723	2053.2	Contig48	p	913	1896	+	-	98%		+
SEQ ID No. 2379	457.2	Contig48	p	2184	3791		-	94%		-
SEQ ID No. 2362	454.1	Contig48	p	3988	5505	-	-	93%		-
SEQ ID No. 2354	453.3	Contig48	p	5528	6304	-	-	98%	+	-
SEQ ID No. 1603	3417.1	Contig48	p	6528	7109	-	-	100%		-
SEQ ID No. 1602	3416.1	Contig48	m	7125	7490	-	-	100%		-
SEQ ID No. 1601	3415.2	Contig48	m	7689	8420	-	-	97%		-
SEQ ID No. 1782	3671.2	Contig48	m	8410	9171	-	-	96%		-
SEQ ID No. 1781	3670.2	Contig48	p	9712	10764	-	-	97%		+
SEQ ID No. 1780	3669.1	Contig48	m	10859	11821	-	-	99%		-
SEQ ID No. 1779	3667.1	Contig48	p	12363	12653	-	-	100%		+
SEQ ID No. 1778	3666.2	Contig48	p	12768	14036	-	+			-
SEQ ID No. 1777	3665.1	Contig48	p	14015	14953	-	+			-
SEQ ID No. 1776	3664.1	Contig48	p	15069	16226	-	+			-
SEQ ID No. 1775	3663.1	Contig48	p	16350	17441	-	+			+
SEQ ID No. 900	2351.2	Contig48	p	17599	18432	-	-	94%		-
SEQ ID No. 899	2350.2	Contig48	m	17731	18108		-	83%		+
SEQ ID No. 901	2352.2	Contig48	p	18433	19479	-	-	88%		-
SEQ ID No. 902	2353.3	Contig48	m	19451	20626	-	-	95%		-
SEQ ID No. 2153	4247.2	Contig48	p	20746	21615		-	76%		+

SEQ ID No. 111	1080.2	Contig48	p	21518	22444	-	+			-
SEQ ID No. 112	1081.2	Contig48	m	22650	23873	-	+			+
SEQ ID No. 2154	4248.1	Contig48	p	24264	24533		+			-
SEQ ID No. 2155	4249.1	Contig48	p	24436	25785	-	-	96%		-
SEQ ID No. 2157	4250.1	Contig48	m	25684	26193		+			+
SEQ ID No. 2158	4251.2	Contig48	m	26261	27592		+			+
SEQ ID No. 639	1891.2	Contig48	m	27597	28370		+			-
SEQ ID No. 756	2112.2	Contig48	p	28628	30067		+			-
SEQ ID No. 3337	838.3	Contig48	m	30280	31686	-	-	99%	+	-
SEQ ID No. 3339	840.2	Contig48	p	31793	32653	-	-	95%		-
SEQ ID No. 3340	841.3	Contig48	p	32857	33192	-	-	100%		-
SEQ ID No. 806	2193.4	Contig48	p	33494	34777	-	-	100%		-
SEQ ID No. 807	2195.1	Contig48	p	34920	36389	-	-	99%		-
SEQ ID No. 808	2196.1	Contig48	p	36362	37102	-	-	99%		-
SEQ ID No. 809	2197.1	Contig48	p	37090	37314	-	+			+
SEQ ID No. 810	2198.1	Contig48	m	37352	37684	-	-	99%		-
SEQ ID No. 811	2199.2	Contig48	m	37725	38468	-	-	98%		-
SEQ ID No. 1197	2799.1	Contig48	m	38556	38960	-	-	99%	+	-
SEQ ID No. 1198	2800.1	Contig48	p	39290	40063	-	-	97%	+	-
SEQ ID No. 689	1986.2	Contig48	p	40082	41437	-	-	97%		+
SEQ ID No. 688	1983.1	Contig48	p	41412	42260	-	-	99%		-
SEQ ID No. 2035	408.3	Contig48	p	42286	43932	+	-	99%		-
SEQ ID No. 2042	409.2	Contig48	p	43914	45074	-	-	98%		-
SEQ ID No. 2048	410.1	Contig48	p	45028	45756	-	-	99%		-
SEQ ID No. 2063	412.3	Contig48	p	45684	46946	+	-	98%		-
SEQ ID No. 919	2373.2	Contig48	p	47260	48243	+	-	99%		-
SEQ ID No. 920	2374.2	Contig48	p	48553	49464	-	+			+
SEQ ID No. 590	1815.2	Contig48	p	49732	50319	-	+			-
SEQ ID No. 591	1817.2	Contig48	p	50616	51212	-	-	98%		+
SEQ ID No. 2782	5208.1	Contig48	p	51661	52008	-	+			-
SEQ ID No. 1329	3005.2	Contig48	p	51996	52523	-	+			-
SEQ ID No. 1462	321.3	Contig48	p	53567	54463	-	+			-
SEQ ID No. 1466	322.3	Contig48	p	54569	54796	-	+			+
SEQ ID No. 1472	323.2	Contig48	m	54848	56335	-	-	97%		+
SEQ ID No. 719	2044.2	Contig48	p	57557	59509	-	-	98%		-
SEQ ID No. 523	172.1	Contig48	m	59557	59775	-	+			-

SEQ ID No. 537	174.1	Contig48	p	60382	61275	-	-	96%	-
SEQ ID No. 551	176.1	Contig48	m	61299	62291	-	-	96%	+
SEQ ID No. 566	178.1	Contig48	p	62390	62704	-	-	100%	+
SEQ ID No. 574	179.1	Contig48	m	62747	63022	-	-	95%	-
SEQ ID No. 581	180.1	Contig48	m	63013	63207	+	+		-
SEQ ID No. 594	182.3	Contig48	m	63200	64855	-	-	99%	-
SEQ ID No. 1330	3008.1	Contig48	m	64939	65592	-	-	98%	-
SEQ ID No. 1059	2580.2	Contig48	m	65564	66331	-	-	99%	-
SEQ ID No. 1057	2579.2	Contig48	m	66441	66896	-	-	98%	-
SEQ ID No. 1331	3009.2	Contig48	m	67167	68123	+	+		+
SEQ ID No. 2365	4546.3	Contig48	p	68632	69690	-	+		-
SEQ ID No. 916	2369.3	Contig48	m	69875	70516	+	-	100%	-
SEQ ID No. 917	2371.1	Contig48	m	70539	71579	-	-	97%	-
SEQ ID No. 918	2372.3	Contig48	m	71633	72022	-	-	97%	-
SEQ ID No. 2364	4542.4	Contig48	p	72350	73981	+	-	99%	-
SEQ ID No. 503	1689.3	Contig48	p	74054	75268	+	-	100%	-
SEQ ID No. 505	1691.3	Contig48	p	75250	76881	-	-	99%	-
SEQ ID No. 546	1752.2	Contig48	p	76882	77436	+	-	100%	-
SEQ ID No. 545	1750.3	Contig48	p	77589	78464	-	-	97%	-
SEQ ID No. 1243	2884.1	Contig48	m	78487	78768	-	-	94%	+
SEQ ID No. 152	114.2	Contig48	p	79316	79507	-	+		+
SEQ ID No. 157	115.2	Contig48	p	79533	80096	-	+		-
SEQ ID No. 163	116.1	Contig48	p	80429	80992	-	+		-
SEQ ID No. 3159	6083.1	Contig48	p	81259	81510		+		+
SEQ ID No. 170	117.1	Contig48	p	81479	81859		+		+
SEQ ID No. 176	118.1	Contig48	p	81813	82490	-	+		+
SEQ ID No. 183	119.1	Contig48	p	82812	83327	-	+		-
SEQ ID No. 194	121.1	Contig48	m	83648	84016	-	+		-
SEQ ID No. 3043	5765.1	Contig48	m	84091	84564	-	+		-
SEQ ID No. 199	122.3	Contig48	m	84504	85001	+	-	100%	-
SEQ ID No. 687	1979.1	Contig48	m	85012	85803	-	-	99%	-
SEQ ID No. 686	1978.1	Contig48	m	85785	86507	-	-	97%	-
SEQ ID No. 685	1976.2	Contig48	m	86408	88291	-	-	98%	-
SEQ ID No. 793	2166.2	Contig48	m	88413	90341	-	-	97%	+
SEQ ID No. 794	2167.2	Contig48	m	90554	92077	-	-	99%	-
SEQ ID No. 1242	2882.2	Contig48	p	92029	92766	+	-	99%	-

SEQ ID No. 3044	5766.1	Contig48	p	92800	93363	+	-	98%		-
SEQ ID No. 278	1341.2	Contig48	p	93650	94402	-	-	98%		+
SEQ ID No. 277	1339.2	Contig48	m	94458	95033	-	-	100%		-
SEQ ID No. 276	1338.3	Contig48	p	95203	97308	-	-	98%		-
SEQ ID No. 3282	757.1	Contig48	p	97275	97922	-	-	99%		-
SEQ ID No. 3283	758.1	Contig48	p	97900	98550	-	-	98%	+	-
SEQ ID No. 3284	760.2	Contig48	p	98779	100542	-	-	99%		+
SEQ ID No. 1741	3614.3	Contig48	m	100655	101728	-	-	99%		-
SEQ ID No. 2119	4194.3	Contig48	p	101959	103278	-	-	99%		+
SEQ ID No. 3330	825.4	Contig48	m	103305	104558	+	-	98%		-
SEQ ID No. 2885	5438.1	Contig48	m	104563	105147		+			-
SEQ ID No. 3329	823.4	Contig48	m	105148	105969	-	-	98%		-
SEQ ID No. 3328	821.2	Contig48	p	106064	106960	-	-	98%		+
SEQ ID No. 3158	6081.1	Contig48	m	107058	107270		+			-
SEQ ID No. 1946	396.4	Contig48	m	107186	108082	-	-	98%		-
SEQ ID No. 1934	393.1	Contig48	p	108405	108956	-	-	100%	+	-
SEQ ID No. 3379	897.2	Contig48	p	108957	111251	-	-	99%	+	-
SEQ ID No. 3380	898.2	Contig48	p	111179	112054	-	-	98%		-
SEQ ID No. 3381	899.3	Contig48	p	112066	112842	-	-	98%		-
SEQ ID No. 1814	3720.2	Contig48	p	112836	113633	-	-	97%	+	-
SEQ ID No. 2884	5437.2	Contig48	p	113664	113978	+	-	94%		+
SEQ ID No. 3045	5767.1	Contig48	m	113682	114005		-	93%		+
SEQ ID No. 3046	5769.2	Contig48	m	114096	114503	-	-	99%		-
SEQ ID No. 423	1569.6	Contig48	m	114485	116662	-	-	99%		-
SEQ ID No. 951	2422.3	Contig48	m	116663	117175	-	-	98%	+	-
SEQ ID No. 952	2423.2	Contig48	m	117160	118278	-	-	98%		-
SEQ ID No. 2922	551.2	Contig48	p	118410	119204	-	-	94%		-
SEQ ID No. 2933	553.1	Contig48	p	119189	119620	-	+			+
SEQ ID No. 2946	555.2	Contig48	m	119679	120698	-	-	96%		-
SEQ ID No. 1360	3058.1	Contig48	p	120708	121190	-	-	100%		+
SEQ ID No. 863	2291.2	Contig48	p	121301	122362	-	-	100%		-
SEQ ID No. 862	2290.1	Contig48	p	122363	123283	-	-	98%		-
SEQ ID No. 861	2289.2	Contig48	p	123222	123761	-	-	98%		-
SEQ ID No. 1359	3057.1	Contig48	p	123677	124783	-	-	99%		-
SEQ ID No. 1358	3055.1	Contig48	p	124770	125705		-	98%		-
SEQ ID No. 1357	3054.3	Contig48	p	125624	126583	-	-	92%		-

SEQ ID No. 1356	3053.3	Contig48	m	126062	126499		-	86%		-
SEQ ID No. 3157	6079.1	Contig48	m	126066	126479	-	-	83%		-
SEQ ID No. 736	2070.4	Contig48	p	126562	128229	+	-	98%		-
SEQ ID No. 68	1014.2	Contig48	p	128457	129848	-	-	98%		-
SEQ ID No. 67	1013.1	Contig48	m	128699	129295	-	-	97%		+
SEQ ID No. 734	2068.3	Contig48	p	130078	131586	+	-	98%		+
SEQ ID No. 2448	4669.2	Contig48	p	131759	132040	-	-	97%		-
SEQ ID No. 2447	4668.1	Contig48	m	132141	133445	-	-	98%		-
SEQ ID No. 2446	4666.2	Contig48	p	133569	134588	+	-	98%		-
SEQ ID No. 365	1482.4	Contig48	p	134694	135689	-	-	98%		-
SEQ ID No. 714	2034.2	Contig48	m	136325	137179	-	-	96%		-
SEQ ID No. 2836	5334.1	Contig48	p	137595	138359	-	-	83%		-
SEQ ID No. 1427	315.3	Contig48	p	137887	138900		+			-
SEQ ID No. 1441	317.1	Contig48	m	138934	139956	-	-	99%	+	-
SEQ ID No. 1449	318.2	Contig48	m	139970	141388	-	-	99%	+	-
SEQ ID No. 589	1812.4	Contig48	m	141531	142739	+	-	99%		-
SEQ ID No. 866	2295.3	Contig48	m	142854	143399	-	-	96%		-
SEQ ID No. 867	2297.3	Contig48	m	143330	143803	-	-	99%		-
SEQ ID No. 2611	4907.2	Contig48	m	143912	144691	-	-	98%		-
SEQ ID No. 2837	5337.1	Contig48	m	144919	145224	-	+			-
SEQ ID No. 2838	5338.2	Contig48	m	145238	146323	-	-	100%		-
SEQ ID No. 2165	4266.3	Contig48	p	146551	147867	-	-	99%		-
SEQ ID No. 3258	723.3	Contig48	p	148217	150046	-	-	99%		-
SEQ ID No. 3257	722.3	Contig48	p	150016	150708	-	-	97%		-
SEQ ID No. 3256	719.4	Contig48	p	150832	153201	-	-	99%		-
SEQ ID No. 3011	5671.2	Contig48	p	153153	153584	-	-	100%		-
SEQ ID No. 3009	5669.2	Contig48	p	153548	155221	-	-	99%		+
SEQ ID No. 3008	5668.2	Contig48	m	155278	157158	-	-	97%		+
SEQ ID No. 3005	5656.2	Contig48	p	157395	158135	-	-	99%		-
SEQ ID No. 3092	590.4	Contig48	p	158250	161354	-	+			+
SEQ ID No. 2398	4599.2	Contig48	m	161403	162059	-	+			-
SEQ ID No. 3156	6074.1	Contig49	p	3	614	-	+			-
SEQ ID No. 2278	4421.2	Contig49	p	599	1414	-	+			-
SEQ ID No. 2279	4422.1	Contig49	p	1499	1792		+			+
SEQ ID No. 2280	4424.1	Contig49	p	1861	2124		+			+
SEQ ID No. 2281	4425.1	Contig49	p	2147	2680	-	-	94%		-

SEQ ID No. 2282	4427.1	Contig49	m	3129	3356	-	+		+
SEQ ID No. 2283	4428.1	Contig49	p	3412	3774		+		+
SEQ ID No. 2284	4429.2	Contig49	m	3570	3890		-	91%	+
SEQ ID No. 597	1827.3	Contig49	p	4176	5615	-	+		+
SEQ ID No. 596	1826.2	Contig49	p	5708	6622	-	-	95%	+
SEQ ID No. 2621	4920.4	Contig49	m	6970	7851	-	-	94%	-
SEQ ID No. 256	1302.2	Contig49	m	7878	9074	-	-	99%	-
SEQ ID No. 257	1303.3	Contig49	m	9075	9848	-	-	97%	-
SEQ ID No. 1709	3570.2	Contig49	m	9817	10080	-	+		-
SEQ ID No. 931	2392.2	Contig49	m	10206	10697		-	96%	+
SEQ ID No. 930	2391.2	Contig49	p	10897	11898	-	-	84%	+
SEQ ID No. 1710	3572.1	Contig49	m	12027	12509	-	-	96%	-
SEQ ID No. 428	1578.3	Contig49	p	12817	13644	-	-	92%	-
SEQ ID No. 429	1579.4	Contig49	m	13726	14193	+	-	92%	-
SEQ ID No. 2694	5044.2	Contig49	m	14349	15251	-	-	98%	-
SEQ ID No. 2695	5046.2	Contig49	m	15500	15922	-	-	98%	-
SEQ ID No. 2696	5047.3	Contig49	m	16011	16757	-	-	98%	-
SEQ ID No. 1156	2739.3	Contig49	m	16733	17521	-	-	99%	-
SEQ ID No. 1155	2738.2	Contig49	m	17487	18263		-	99%	+
SEQ ID No. 1154	2737.2	Contig49	m	18508	19089	-	-	93%	+
SEQ ID No. 645	190.3	Contig49	p	19228	19989	-	-	93%	+
SEQ ID No. 651	191.1	Contig49	m	20119	20982	-	-	98%	-
SEQ ID No. 662	193.3	Contig49	p	21045	22139	-	-	97%	-
SEQ ID No. 675	195.3	Contig49	p	22114	22965	-	-	98%	-
SEQ ID No. 683	197.3	Contig49	p	22941	24404	-	-	96%	-
SEQ ID No. 691	199.1	Contig49	p	24389	25549	+	-	96%	-
SEQ ID No. 700	201.2	Contig49	p	25543	26778	-	-	97%	-
SEQ ID No. 712	2030.2	Contig49	m	26936	28129	-	+		+
SEQ ID No. 718	2041.2	Contig49	p	28480	29592	-	-	94%	-
SEQ ID No. 716	2039.1	Contig49	p	29682	30515		+		+
SEQ ID No. 2629	494.2	Contig49	m	30690	31535	+	-	95%	-
SEQ ID No. 2642	496.1	Contig49	p	31594	32652		-	85%	+
SEQ ID No. 2650	497.1	Contig49	m	32699	33202	-	-	90%	+
SEQ ID No. 2655	498.4	Contig49	p	33556	34560	-	-	92%	-
SEQ ID No. 3042	5761.1	Contig49	p	34545	35045	-	+		-
SEQ ID No. 715	2037.3	Contig49	p	35275	36714	+	-	96%	-

SEQ ID No. 1153	2733.1	Contig49	p	36823	37893	-	-	89%		-
SEQ ID No. 1152	2732.1	Contig49	m	38090	38974	+	-	93%		+
SEQ ID No. 1151	2730.1	Contig49	p	39326	40288	-	-	97%		+
SEQ ID No. 182	1188.2	Contig49	m	40523	42844	+	-	98%		-
SEQ ID No. 663	1930.3	Contig49	m	42892	43923	-	-	94%		+
SEQ ID No. 720	2049.2	Contig49	m	44183	45454	-	-	89%		+
SEQ ID No. 1149	2728.1	Contig49	p	45792	46145	-	-	97%		-
SEQ ID No. 1148	2727.2	Contig49	p	46358	49045	-	+			+
SEQ ID No. 197	1213.2	Contig49	p	49373	50596	+	+			-
SEQ ID No. 196	1211.3	Contig49	p	50962	52008		-	86%		-
SEQ ID No. 3404	930.3	Contig49	p	52100	53491	-	-	99%	+	-
SEQ ID No. 3405	932.1	Contig49	p	53437	54480	-	-	98%		-
SEQ ID No. 3406	935.2	Contig49	m	54481	55269	-	-	99%		-
SEQ ID No. 3407	936.3	Contig49	m	55263	55808		-	97%		+
SEQ ID No. 1177	2769.2	Contig49	m	55789	57639	-	-	96%		-
SEQ ID No. 1179	2770.1	Contig49	p	57644	58357	-	-	96%	+	-
SEQ ID No. 1180	2771.2	Contig49	p	58350	59543	-	-	97%		-
SEQ ID No. 3104	5926.1	Contig49	p	59524	59913	-	+			-
SEQ ID No. 1181	2772.1	Contig49	m	59918	60340	-	-	98%		-
SEQ ID No. 3214	657.3	Contig49	p	60518	63580	-	-	96%		-
SEQ ID No. 3213	656.3	Contig49	p	63581	65569	-	-	94%		+
SEQ ID No. 1655	350.3	Contig49	p	65586	66236	-	-	98%		+
SEQ ID No. 1664	351.1	Contig49	m	66378	66824		-	91%		-
SEQ ID No. 1685	354.1	Contig49	m	66885	68309	-	-	97%		-
SEQ ID No. 1708	357.2	Contig49	m	68395	69129	-	-	98%		-
SEQ ID No. 1719	359.2	Contig49	m	69080	69700		-	95%		-
SEQ ID No. 1182	2774.1	Contig49	p	69554	70459	-	-	97%		-
SEQ ID No. 1183	2775.1	Contig49	p	70567	70794	-	+			-
SEQ ID No. 1184	2777.1	Contig49	m	70880	71971	-	-	94%		-
SEQ ID No. 1185	2778.2	Contig49	m	71956	72480	-	-	97%		-
SEQ ID No. 1187	2780.2	Contig49	m	72516	73382	-	-	97%		-
SEQ ID No. 1188	2781.1	Contig49	m	73383	73757	-	-	99%		-
SEQ ID No. 478	1649.3	Contig49	m	73750	76110	-	-	94%		-
SEQ ID No. 479	1650.2	Contig49	m	76053	77585		-	96%		-
SEQ ID No. 1063	2587.2	Contig49	p	77654	78526	-	-	96%		-
SEQ ID No. 669	1940.2	Contig49	p	78565	80157	+	-	98%		-

SEQ ID No. 670	1943.3	Contig49	p	80136	81140	+	-	98%	+	-
SEQ ID No. 2375	4561.2	Contig49	p	81131	82534	-	-	98%		-
SEQ ID No. 2374	4560.2	Contig49	p	82681	83823	-	-	92%		+
SEQ ID No. 2373	4559.1	Contig49	m	83994	85562	+	-	98%		-
SEQ ID No. 2372	4558.1	Contig49	m	85707	86597	-	-	97%		+
SEQ ID No. 3155	6072.1	Contig49	p	86701	86892		+			-
SEQ ID No. 3154	6071.1	Contig49	p	86822	87034		+			-
SEQ ID No. 2371	4554.2	Contig49	m	87282	87542		+			-
SEQ ID No. 2370	4553.2	Contig49	p	87651	88352	+	+			-
SEQ ID No. 2752	5146.2	Contig49	p	88433	89740	+	+			-
SEQ ID No. 1402	3116.2	Contig49	m	89715	90125	-	+			-
SEQ ID No. 1401	3114.2	Contig49	p	90204	91163		+			-
SEQ ID No. 1400	3113.1	Contig49	p	91157	93160	-	+			-
SEQ ID No. 1399	3110.1	Contig49	p	93148	93756	-	+			-
SEQ ID No. 1398	3109.2	Contig49	p	93743	94096	-	+			-
SEQ ID No. 1397	3107.2	Contig49	m	94191	94412		+			+
SEQ ID No. 1396	3106.2	Contig49	m	94616	96460	-	-	96%		-
SEQ ID No. 1395	3105.2	Contig49	m	96450	96755		+			-
SEQ ID No. 1394	3104.2	Contig49	m	96892	99459	-	-	97%		-
SEQ ID No. 1393	3102.2	Contig49	p	99671	100228	-	-	98%		-
SEQ ID No. 1392	3101.1	Contig49	p	100453	102348	-	+			-
SEQ ID No. 2751	514.5	Contig49	m	102459	107018	-	-	95%		+
SEQ ID No. 2235	436.4	Contig49	p	104371	104940		-	96%		+
SEQ ID No. 783	2150.1	Contig49	m	107367	107894	-	-	97%		-
SEQ ID No. 201	1222.5	Contig49	m	107944	109923	-	-	96%		+
SEQ ID No. 877	2315.3	Contig49	p	110157	110999	-	-	93%		-
SEQ ID No. 878	2317.2	Contig49	p	111016	112482	-	-	97%		-
SEQ ID No. 1199	2804.1	Contig49	p	112445	113809	-	-	98%		-
SEQ ID No. 858	2279.2	Contig49	p	113883	114656	-	+			+
SEQ ID No. 525	1722.2	Contig49	p	114908	115852	-	-	97%		-
SEQ ID No. 526	1723.5	Contig49	p	115955	116539	-	-	98%		-
SEQ ID No. 3153	6070.1	Contig49	p	116472	116879	-	+			-
SEQ ID No. 1200	2806.1	Contig49	m	117016	117618	-	-	99%		-
SEQ ID No. 3223	670.1	Contig49	m	117945	119357	-	-	97%		-
SEQ ID No. 1385	309.2	Contig49	m	119391	122780	-	-	98%		-
SEQ ID No. 1391	310.1	Contig49	p	122902	123675	-	-	96%		+

SEQ ID No. 1406	312.3	Contig49	p	123904	124869	-	-	98%	+
SEQ ID No. 1201	2807.1	Contig49	p	124936	125370	-	-	95%	-
SEQ ID No. 962	2434.2	Contig49	m	125457	126668	-	-	97%	-
SEQ ID No. 961	2432.2	Contig49	m	126730	127992	-	-	97%	-
SEQ ID No. 1202	2808.1	Contig49	p	128188	128580	-	-	98%	-
SEQ ID No. 953	2424.3	Contig49	p	128735	129358	-	-	97%	-
SEQ ID No. 954	2425.2	Contig49	p	129470	130117	-	-	97%	-
SEQ ID No. 955	2426.2	Contig49	p	130209	130778	-	-	94%	+
SEQ ID No. 1203	2810.1	Contig49	p	130891	132279	+	-	98%	-
SEQ ID No. 1204	2811.2	Contig49	p	132257	133306	-	-	98%	+
SEQ ID No. 1764	365.4	Contig49	p	133354	134355	-	-	98%	-
SEQ ID No. 1758	364.2	Contig49	p	134520	136475	-	-	98%	-
SEQ ID No. 1730	360.3	Contig49	p	136479	138278	-	-	97%	-
SEQ ID No. 713	2031.2	Contig49	p	138614	138955		+		-
SEQ ID No. 1436	3163.1	Contig49	p	138969	139274	-	+		-
SEQ ID No. 1437	3165.1	Contig49	p	139471	139872	-	-	98%	+
SEQ ID No. 1438	3166.1	Contig49	m	139958	140611	-	-	96%	-
SEQ ID No. 1439	3167.1	Contig49	p	140837	142129	-	-	100%	-
SEQ ID No. 1440	3169.1	Contig49	p	142184	142906		-	98%	-
SEQ ID No. 1000	2488.2	Contig49	m	143043	144032		+		+
SEQ ID No. 1001	2489.2	Contig49	m	144070	144399		+		+
SEQ ID No. 1003	2490.2	Contig49	p	144461	144931		+		+
SEQ ID No. 2483	4719.2	Contig49	m	145036	146628	+	-	95%	-
SEQ ID No. 2482	4718.1	Contig49	p	146787	147245	-	-	96%	+
SEQ ID No. 2481	4716.2	Contig49	m	147260	148393	-	-	96%	-
SEQ ID No. 1128	2695.2	Contig49	p	148478	149104	-	-	93%	-
SEQ ID No. 1129	2696.2	Contig49	p	149173	150738	-	-	95%	-
SEQ ID No. 420	1564.3	Contig49	p	150804	152222	-	-	99%	-
SEQ ID No. 419	1563.3	Contig49	p	152206	152847	-	-	95%	-
SEQ ID No. 1130	2698.1	Contig49	m	152930	153349	-	-	98%	-
SEQ ID No. 1133	2700.2	Contig49	m	153313	154554		-	97%	-
SEQ ID No. 1134	2702.2	Contig49	p	154618	155514	-	-	97%	-
SEQ ID No. 1496	326.2	Contig49	p	155911	159276	-	-	96%	+
SEQ ID No. 1490	325.1	Contig49	p	159347	160909	-	-	94%	+
SEQ ID No. 1482	324.3	Contig49	m	160975	161676	-	+		-
SEQ ID No. 999	2486.3	Contig49	m	161677	163272	-	-	97%	-

SEQ ID No. 1135	2704.4	Contig49	m	163203	164651	-	-	97%	-
SEQ ID No. 1136	2705.3	Contig49	m	164641	165852	-	-	98%	-
SEQ ID No. 1137	2706.2	Contig49	m	165977	166876	-	-	94%	-
SEQ ID No. 179	1183.4	Contig49	p	166928	168976	-	+		+
SEQ ID No. 180	1184.2	Contig49	m	168987	170183	+	-	95%	-
SEQ ID No. 181	1186.2	Contig49	m	170239	171039	-	-	88%	-
SEQ ID No. 638	189.3	Contig49	p	171178	172434	-	-	97%	-
SEQ ID No. 630	188.1	Contig49	m	172539	173675	-	-	96%	+
SEQ ID No. 624	187.1	Contig49	p	173879	174583	-	-	94%	+
SEQ ID No. 612	185.1	Contig49	p	174610	175461	+	-	98%	-
SEQ ID No. 603	184.2	Contig49	p	175678	175977		+		+
SEQ ID No. 682	1968.1	Contig49	p	176650	176925	+	-	97%	-
SEQ ID No. 681	1966.2	Contig49	m	176980	177600	-	-	98%	-
SEQ ID No. 1138	2709.1	Contig49	p	177567	178601	-	-	98%	-
SEQ ID No. 2166	4267.5	Contig49	m	178670	180760	-	-	99%	-
SEQ ID No. 359	1474.4	Contig49	m	180783	181997	-	-	99%	-
SEQ ID No. 358	1473.4	Contig49	m	182064	182630	-	-	98%	-
SEQ ID No. 2167	4269.3	Contig49	m	182594	183289	-	-	95%	+
SEQ ID No. 650	1908.4	Contig49	p	183369	185828	-	-	99%	-
SEQ ID No. 2169	4272.3	Contig49	p	185939	186634	-	-	98%	+
SEQ ID No. 2204	4323.1	Contig49	m	186671	187360	-	-	98%	-
SEQ ID No. 2203	4322.1	Contig49	m	187330	188436	-	-	96%	-
SEQ ID No. 2202	4321.2	Contig49	p	188394	191069	-	-	99%	-
SEQ ID No. 2201	4320.1	Contig49	p	191146	192534	+	-	99%	-
SEQ ID No. 224	1251.2	Contig49	p	192522	193505	-	-	97%	-
SEQ ID No. 223	1250.1	Contig49	p	193501	193998	-	-	93%	-
SEQ ID No. 221	1249.2	Contig49	m	194005	194559		+		-
SEQ ID No. 3184	6134.1	Contig5	m	112	720	-	+		-
SEQ ID No. 3183	6133.1	Contig5	m	943	1218	-	+		-
SEQ ID No. 3182	6132.1	Contig5	m	1232	1546	-	+		-
SEQ ID No. 3152	6066.1	Contig50	p	3	491		+		-
SEQ ID No. 2960	5576.3	Contig50	p	879	1850	-	+		-
SEQ ID No. 2147	4239.4	Contig50	p	1918	2205		+		+
SEQ ID No. 2146	4238.1	Contig50	m	2294	2842	-	+		+
SEQ ID No. 2145	4236.3	Contig50	m	2778	4787	-	+		-
SEQ ID No. 2144	4234.3	Contig50	m	4788	5267		+		+

SEQ ID No. 2143	4233.3	Contig50	p	5113	5352		+		+
SEQ ID No. 459	1624.4	Contig50	m	5355	7022	-	+		+
SEQ ID No. 458	1623.4	Contig50	m	6937	8328	-	+		-
SEQ ID No. 2561	4822.3	Contig50	m	8329	9435	-	+		-
SEQ ID No. 247	1287.3	Contig50	m	9299	12076	-	+		-
SEQ ID No. 2925	5517.1	Contig50	m	12084	12476	+	+		-
SEQ ID No. 2050	4104.2	Contig50	m	12430	13818	-	+		-
SEQ ID No. 2051	4105.1	Contig50	m	13740	14561	+	+		-
SEQ ID No. 2052	4106.1	Contig50	m	14509	15180	+	+		+
SEQ ID No. 2053	4107.1	Contig50	m	15176	15646	-	+		+
SEQ ID No. 914	2367.2	Contig50	m	15535	15933	+	+		+
SEQ ID No. 297	1372.2	Contig50	m	15924	16229	-	+		+
SEQ ID No. 298	1373.1	Contig50	m	16190	16525	-	+		+
SEQ ID No. 299	1375.2	Contig50	m	16526	17281	+	+		-
SEQ ID No. 300	1376.2	Contig50	m	17215	17766	-	+		-
SEQ ID No. 2054	4108.1	Contig50	m	17675	18538	+	+		-
SEQ ID No. 2055	4109.2	Contig50	m	18522	19115	+	+		+
SEQ ID No. 3060	5814.1	Contig50	m	19109	19555	+	+		-
SEQ ID No. 3059	5813.1	Contig50	m	19540	19761	-	+		-
SEQ ID No. 3058	5812.1	Contig50	m	19757	20260	-	+		+
SEQ ID No. 1086	2625.2	Contig50	m	20085	21020	-	+		+
SEQ ID No. 1087	2626.1	Contig50	p	21133	21804	-	+		-
SEQ ID No. 583	1803.2	Contig50	p	21826	22089		+		-
SEQ ID No. 584	1804.2	Contig50	p	22255	23502	+	-	97%	-
SEQ ID No. 1088	2627.1	Contig50	p	23499	24755	+	-	98%	-
SEQ ID No. 1090	2631.4	Contig50	p	24756	27710	-	-	99%	-
SEQ ID No. 3112	5953.1	Contig50	p	27632	27916	-	+		-
SEQ ID No. 3273	747.3	Contig50	p	27917	30127	-	-	99%	-
SEQ ID No. 2761	5164.1	Contig50	p	30361	30561	+	+		+
SEQ ID No. 3272	744.3	Contig50	p	30803	31477	-	+		-
SEQ ID No. 3271	743.3	Contig50	m	31500	31961	-	-	94%	-
SEQ ID No. 3111	5952.1	Contig50	p	32298	32483		+		-
SEQ ID No. 3425	96.6	Contig50	m	32725	33612	-	-	99%	-
SEQ ID No. 3432	97.2	Contig50	m	33710	33985		-	97%	+
SEQ ID No. 3446	99.1	Contig50	m	34081	34362		-	94%	+
SEQ ID No. 2760	5162.1	Contig50	p	34102	34254	-	-	100%	+

SEQ ID No. 69	102.1	Contig50	p	34385	37093	-	-	96%	-
SEQ ID No. 75	103.1	Contig50	p	37068	38339	+	-	98%	-
SEQ ID No. 83	104.1	Contig50	p	38294	39469	+	-	98%	-
SEQ ID No. 103	107.1	Contig50	p	39535	42687	-	-	98%	-
SEQ ID No. 117	109.1	Contig50	p	42543	44045	-	-	94%	-
SEQ ID No. 133	111.2	Contig50	p	43985	45556	-	-	96%	-
SEQ ID No. 851	227.2	Contig50	p	45523	46464	-	-	93%	-
SEQ ID No. 859	228.1	Contig50	p	46477	46656	-	-	97%	+
SEQ ID No. 892	234.2	Contig50	p	46705	48930	-	-	99%	-
SEQ ID No. 680	1961.3	Contig50	p	49060	49950	-	-	96%	+
SEQ ID No. 390	152.3	Contig50	p	50084	50461	+	-	98%	-
SEQ ID No. 398	153.2	Contig50	p	50462	51799	+	-	99%	-
SEQ ID No. 405	154.1	Contig50	p	51796	52764	+	-	96%	-
SEQ ID No. 424	157.2	Contig50	p	52765	55980	+	-	98%	-
SEQ ID No. 3107	594.1	Contig50	p	55956	56429	+	-	96%	+
SEQ ID No. 3105	593.2	Contig50	m	56568	56918	-	-	96%	+
SEQ ID No. 3102	592.3	Contig50	p	57113	58795	-	-	94%	-
SEQ ID No. 504	169.1	Contig50	m	58808	59020	-	+		+
SEQ ID No. 497	168.1	Contig50	p	59115	59687	-	-	97%	-
SEQ ID No. 486	166.2	Contig50	p	59537	62800	-	-	95%	-
SEQ ID No. 464	163.1	Contig50	p	62978	64276	-	+		-
SEQ ID No. 447	160.1	Contig50	m	64504	64863		+		-
SEQ ID No. 438	159.1	Contig50	p	65005	65316		-	89%	-
SEQ ID No. 430	158.2	Contig50	p	65304	65987	-	-	98%	-
SEQ ID No. 679	1960.2	Contig50	p	65972	66484	+	-	92%	+
SEQ ID No. 462	1628.3	Contig50	m	66550	67515	+	-	92%	-
SEQ ID No. 463	1629.1	Contig50	m	67365	68138	+	-	95%	-
SEQ ID No. 465	1631.3	Contig50	m	68126	70117	-	-	97%	-
SEQ ID No. 1091	2637.1	Contig50	p	70315	70731	-	-	100%	+
SEQ ID No. 3325	818.2	Contig50	m	70763	71647	-	-	94%	+
SEQ ID No. 3326	819.2	Contig50	m	71690	73171	-	-	100%	-
SEQ ID No. 411	1550.2	Contig50	m	73149	73937	-	-	97%	+
SEQ ID No. 410	1549.1	Contig50	m	73898	74191	+	-	100%	-
SEQ ID No. 409	1548.2	Contig50	m	74170	74871	-	-	99%	-
SEQ ID No. 336	1425.2	Contig50	m	74865	75158	-	-	97%	-
SEQ ID No. 335	1424.2	Contig50	m	75133	75561	-	-	96%	-

SEQ ID No. 334	1423.2	Contig50	m	75552	76979	-	-	97%		-
SEQ ID No. 1092	2639.1	Contig50	m	77314	77775	+	-	95%		-
SEQ ID No. 1470	3226.2	Contig50	m	78822	79085	-	-	97%		-
SEQ ID No. 708	2024.2	Contig50	p	79435	79977	-	-	98%		-
SEQ ID No. 363	148.2	Contig50	p	79937	81463	-	-	96%		-
SEQ ID No. 376	150.2	Contig50	p	81725	82705	-	+			-
SEQ ID No. 2902	5473.1	Contig50	p	82659	83573	-	+			-
SEQ ID No. 382	151.2	Contig50	m	83861	84685	-	-	97%		-
SEQ ID No. 1471	3228.1	Contig50	m	84690	84905	-	+			-
SEQ ID No. 1473	3230.1	Contig50	m	84994	85458		-	90%		+
SEQ ID No. 1474	3231.1	Contig50	m	85582	85800	-	+			-
SEQ ID No. 1475	3232.1	Contig50	m	85971	86342	-	+			+
SEQ ID No. 1476	3233.1	Contig50	m	86294	87133	-	+			-
SEQ ID No. 1477	3234.3	Contig50	m	87124	87846	-	+			-
SEQ ID No. 2903	5474.2	Contig50	p	87816	88361	-	+			-
SEQ ID No. 3363	875.5	Contig50	p	88330	89613	+	+			-
SEQ ID No. 3364	876.1	Contig50	p	89606	90100	-	+			-
SEQ ID No. 3365	879.2	Contig50	p	90587	91051	-	+			-
SEQ ID No. 2982	5615.1	Contig50	p	91012	91698	-	+			-
SEQ ID No. 3103	5920.2	Contig50	p	91731	92939	-	-	97%	+	-
SEQ ID No. 3245	704.4	Contig50	m	93538	94440	-	-	98%		+
SEQ ID No. 3246	705.1	Contig50	m	94413	95726	-	-	99%		-
SEQ ID No. 3247	706.2	Contig50	p	95685	96212		-	97%		-
SEQ ID No. 3248	707.2	Contig50	m	95702	96757	-	-	97%		-
SEQ ID No. 693	1992.2	Contig50	p	96880	97956	-	-	96%		+
SEQ ID No. 1637	3472.1	Contig50	m	98100	98402		-	97%		-
SEQ ID No. 1638	3473.1	Contig50	m	98526	99380	-	-	92%		+
SEQ ID No. 1639	3476.1	Contig50	p	99459	100595	-	-	98%		-
SEQ ID No. 1640	3477.2	Contig50	p	100596	101333	-	-	99%		-
SEQ ID No. 1753	3633.3	Contig50	p	101324	102286	+	-	98%		-
SEQ ID No. 1754	3634.1	Contig50	p	102287	102892	-	-	100%		-
SEQ ID No. 1755	3636.1	Contig50	p	102978	103520	+	-	100%		-
SEQ ID No. 1756	3637.1	Contig50	p	103508	104485	+	-	98%		-
SEQ ID No. 1757	3638.2	Contig50	p	104476	105135	-	-	97%		-
SEQ ID No. 1759	3640.2	Contig50	p	105120	106100	-	-	99%		-
SEQ ID No. 673	1947.3	Contig50	p	106081	106968	-	-	98%		-

SEQ ID No. 674	1949.4	Contig50	m	106981	107274	-	-	100%		-
SEQ ID No. 828	2231.3	Contig50	m	107472	108779	-	-	99%	+	-
SEQ ID No. 827	2230.2	Contig50	m	108772	109404	-	-	99%		-
SEQ ID No. 1760	3641.1	Contig50	m	109367	110134	-	-	99%		-
SEQ ID No. 1761	3642.2	Contig50	m	110241	111392	-	-	99%		-
SEQ ID No. 2724	5087.2	Contig50	p	111723	112805	-	+			+
SEQ ID No. 270	1330.5	Contig50	m	113327	115459	-	-	98%		-
SEQ ID No. 269	1328.2	Contig50	p	115459	116604	-	-	100%		-
SEQ ID No. 268	1325.2	Contig50	p	116708	118477	-	-	98%		-
SEQ ID No. 167	1163.2	Contig50	p	118435	119172	-	-	99%		-
SEQ ID No. 166	1162.1	Contig50	m	119212	119568	-	-	98%		-
SEQ ID No. 165	1161.2	Contig50	p	119787	120848	-	-	98%	+	-
SEQ ID No. 1897	3867.2	Contig50	m	121146	122486	-	-	99%		-
SEQ ID No. 1896	3863.2	Contig50	p	122748	123362	-	-	98%		-
SEQ ID No. 1976	3999.3	Contig50	p	123480	125441	-	-	99%		-
SEQ ID No. 1975	3998.3	Contig50	m	125025	125495		-	96%		-
SEQ ID No. 3357	866.2	Contig50	p	125624	126769	-	-	100%		-
SEQ ID No. 3356	865.2	Contig50	p	126853	127977	-	-	98%		-
SEQ ID No. 460	1625.3	Contig50	p	128135	129334	-	-	99%		-
SEQ ID No. 461	1626.3	Contig50	p	129335	130678	-	-	99%		-
SEQ ID No. 2299	4457.2	Contig50	p	130769	131809	-	-	97%		-
SEQ ID No. 3030	572.2	Contig50	p	131938	133629	+	-	98%		-
SEQ ID No. 3034	573.1	Contig50	m	133783	134067	-	-	100%		-
SEQ ID No. 3038	574.2	Contig50	m	134078	134701	-	-	99%		+
SEQ ID No. 2806	5268.2	Contig50	m	134740	135327		-	98%		-
SEQ ID No. 1405	3119.1	Contig50	m	135659	136453	-	-	94%		-
SEQ ID No. 1404	3118.1	Contig50	p	135939	136310		-	96%		+
SEQ ID No. 1407	3121.2	Contig50	m	136447	137154	-	-	95%		-
SEQ ID No. 1408	3124.2	Contig50	p	137300	138337	-	-	99%	+	-
SEQ ID No. 1409	3127.1	Contig50	m	138666	139430	-	-	100%	+	-
SEQ ID No. 2491	473.2	Contig50	p	139666	140541	-	-	97%		-
SEQ ID No. 2484	472.1	Contig50	p	140538	141176	-	-	98%		-
SEQ ID No. 2467	470.3	Contig50	p	141191	141430	-	+			-
SEQ ID No. 2458	469.3	Contig50	p	141568	143721	-	-	99%		-
SEQ ID No. 2454	468.3	Contig50	p	143679	144107	-	-	98%		-
SEQ ID No. 1410	3128.1	Contig50	p	144071	144601	-	-	99%		-

SEQ ID No. 1015	2510.2	Contig50	p	144653	146515	-	-	97%		-
SEQ ID No. 1013	2509.3	Contig50	p	146573	147010		-	98%		+
SEQ ID No. 1411	3129.2	Contig50	p	147083	148138	-	-	99%		-
SEQ ID No. 1412	3130.1	Contig50	p	148182	149018	-	-	97%		-
SEQ ID No. 1413	3131.1	Contig50	p	149155	149499	+	-	100%		-
SEQ ID No. 3418	949.3	Contig50	p	149500	151377	+	-	99%	+	-
SEQ ID No. 780	2148.1	Contig50	p	151378	152307	-	-	100%	+	-
SEQ ID No. 779	2147.1	Contig50	m	152485	152835	-	-	100%		-
SEQ ID No. 778	2146.2	Contig50	m	152863	154005	-	-	98%		-
SEQ ID No. 1414	3132.1	Contig50	p	154060	154446		-	95%		+
SEQ ID No. 1415	3133.2	Contig50	p	154385	155191	-	-	97%		+
SEQ ID No. 2807	5269.1	Contig50	p	155067	155423		-	94%		+
SEQ ID No. 2808	5270.2	Contig50	p	155164	155745		-	95%		+
SEQ ID No. 2244	4372.3	Contig50	p	155848	157050	+	-	98%		-
SEQ ID No. 2243	4371.1	Contig50	m	157168	158067		-	99%		-
SEQ ID No. 2242	4370.1	Contig50	p	158463	159020	-	-	98%		+
SEQ ID No. 2241	4369.1	Contig50	p	159079	159750		-	88%		-
SEQ ID No. 2240	4367.1	Contig50	p	159936	160592	-	-	82%		+
SEQ ID No. 2239	4366.1	Contig50	m	160879	161298	-	-	91%		-
SEQ ID No. 2238	4365.4	Contig50	m	161394	161951	-	-	86%		-
SEQ ID No. 2974	5602.3	Contig50	p	162005	162481	-	-	97%		-
SEQ ID No. 3320	808.4	Contig50	p	162640	165525	-	-	78%		+
SEQ ID No. 3319	807.2	Contig50	p	165692	167020		-	93%		-
SEQ ID No. 3318	806.2	Contig50	m	167060	167803	-	-	97%		-
SEQ ID No. 2037	4082.2	Contig50	m	167817	169130	-	-	92%		-
SEQ ID No. 529	1726.4	Contig50	m	169200	170873	-	-	93%		-
SEQ ID No. 2911	5484.4	Contig50	m	170954	172321		-	94%		+
SEQ ID No. 2378	4567.2	Contig50	p	172530	174488	-	-	91%		-
SEQ ID No. 2377	4566.1	Contig50	m	174558	175169	-	-	96%		-
SEQ ID No. 3094	5901.3	Contig50	m	175331	176827	-	-	82%		-
SEQ ID No. 3093	5900.3	Contig50	p	176579	176953		-	84%		+
SEQ ID No. 2444	4663.3	Contig50	m	177123	178034	+	-	97%		-
SEQ ID No. 2443	4662.1	Contig50	m	178485	178745	-	+			+
SEQ ID No. 2442	4660.2	Contig50	m	179180	179395		+			+
SEQ ID No. 2441	4659.2	Contig50	p	179570	179950		-	93%		+
SEQ ID No. 2440	4658.2	Contig50	p	180138	180596	-	-	96%		-

SEQ ID No. 2439	4657.1	Contig50	p	180604	181254	-	-	91%	-
SEQ ID No. 2438	4656.2	Contig50	p	181581	183482	-	-	89%	+
SEQ ID No. 3151	6064.1	Contig50	p	183848	184141		-	92%	+
SEQ ID No. 2842	5344.2	Contig50	m	184198	184968	-	-	90%	-
SEQ ID No. 2841	5342.2	Contig50	m	185006	185581	-	-	94%	-
SEQ ID No. 2840	5341.1	Contig50	p	185774	186055	-	-	96%	+
SEQ ID No. 2839	5340.1	Contig50	p	186391	187947	-	-	81%	-
SEQ ID No. 2657	4984.2	Contig50	m	188128	188745	-	-	97%	-
SEQ ID No. 3150	6062.1	Contig50	p	188912	189430	-	+		-
SEQ ID No. 1533	3314.2	Contig51	p	46	753	-	+		+
SEQ ID No. 1534	3316.1	Contig51	m	1154	2896	-	-	96%	+
SEQ ID No. 726	2057.3	Contig51	p	3624	4694	-	-	99%	-
SEQ ID No. 3437	977.3	Contig51	p	4604	5275		-	94%	+
SEQ ID No. 3436	976.3	Contig51	p	5276	5569	-	-	94%	-
SEQ ID No. 3435	974.2	Contig51	p	5479	6735	-	-	98%	-
SEQ ID No. 725	2056.1	Contig51	p	6907	7809	-	-	96%	-
SEQ ID No. 234	1270.2	Contig51	m	7842	8312	-	-	99%	-
SEQ ID No. 235	1271.3	Contig51	p	8263	8610	-	-	100%	-
SEQ ID No. 3119	5983.1	Contig51	p	8510	8743	-	+		+
SEQ ID No. 236	1272.4	Contig51	p	8710	9696	-	-	97%	-
SEQ ID No. 724	2054.2	Contig51	p	9882	10940	-	-	97%	-
SEQ ID No. 2898	5467.1	Contig51	p	11030	11305		-	95%	+
SEQ ID No. 1535	3319.1	Contig51	m	11064	12434	-	-	96%	+
SEQ ID No. 2645	4962.2	Contig51	m	13063	13476	-	-	96%	+
SEQ ID No. 2644	4961.1	Contig51	m	13669	14061	-	-	96%	-
SEQ ID No. 2643	4960.2	Contig51	m	14167	15003	-	-	96%	+
SEQ ID No. 2632	4943.2	Contig51	m	15275	15778	-	-	92%	+
SEQ ID No. 2631	4942.1	Contig51	m	15834	16097		+		+
SEQ ID No. 2630	4941.2	Contig51	m	16060	16500	-	-	97%	+
SEQ ID No. 775	2141.2	Contig51	m	16809	16967		+		+
SEQ ID No. 3274	748.3	Contig51	p	17140	18522	-	-	98%	-
SEQ ID No. 3275	749.1	Contig51	p	18871	19188	-	-	94%	+
SEQ ID No. 3276	750.2	Contig51	m	19419	19598	-	+		+
SEQ ID No. 3277	751.2	Contig51	p	19507	19821		-	93%	+
SEQ ID No. 3278	752.2	Contig51	m	19664	19837	-	-	92%	+
SEQ ID No. 3149	6061.1	Contig51	p	19993	20172		-	94%	+

SEQ ID No. 776	2143.5	Contig51	m	20064	21584		-	96%		-
SEQ ID No. 871	2302.2	Contig51	m	21648	21992	-	-	98%		-
SEQ ID No. 178	1182.3	Contig51	m	21993	22934	-	-	99%		-
SEQ ID No. 177	1180.2	Contig51	m	22973	23767	-	-	98%		-
SEQ ID No. 175	1179.1	Contig51	m	23986	24621	+	-	83%		+
SEQ ID No. 1140	2712.1	Contig51	m	25013	25828		+			-
SEQ ID No. 1141	2713.2	Contig51	m	25993	27000		+			-
SEQ ID No. 1142	2717.1	Contig51	m	27435	29141	-	-	97%		+
SEQ ID No. 1143	2718.1	Contig51	p	29268	30647	-	-	98%	+	-
SEQ ID No. 992	2477.1	Contig51	p	30907	31392	-	-	98%		-
SEQ ID No. 993	2479.2	Contig51	m	31862	32737	-	-	95%		+
SEQ ID No. 267	1324.4	Contig51	p	33284	34675	-	-	91%		-
SEQ ID No. 737	2073.4	Contig51	m	34700	35431	-	-	98%		-
SEQ ID No. 305	1386.3	Contig51	m	35892	36275	-	-	95%		-
SEQ ID No. 304	1385.2	Contig51	m	36572	37336	-	-	90%		+
SEQ ID No. 738	2075.1	Contig51	m	37420	37815	-	-	100%		+
SEQ ID No. 185	1193.2	Contig51	m	37917	38414		-	98%		+
SEQ ID No. 186	1194.3	Contig51	m	38526	39716	-	-	100%	+	-
SEQ ID No. 1144	2721.3	Contig51	m	39661	41946	-	-	98%	+	-
SEQ ID No. 1145	2724.3	Contig51	p	42127	43302	-	-	97%		+
SEQ ID No. 1146	2725.1	Contig51	p	43206	43565	-	-	100%		-
SEQ ID No. 1147	2726.1	Contig51	p	43667	43867	-	+			+
SEQ ID No. 323	141.4	Contig51	p	44028	46769	-	-	96%		+
SEQ ID No. 331	142.2	Contig51	p	46762	47526	-	-	97%		-
SEQ ID No. 341	144.3	Contig51	p	47728	48822	+	-	98%		-
SEQ ID No. 303	1380.3	Contig51	p	49013	50593	-	-	97%		-
SEQ ID No. 1615	3437.1	Contig51	m	50610	51767		-	96%		+
SEQ ID No. 1617	3442.3	Contig51	m	52578	54371	-	-	99%		-
SEQ ID No. 1618	3443.1	Contig51	m	54463	55263	-	-	98%		-
SEQ ID No. 1619	3444.1	Contig51	m	55371	55877	-	-	98%		+
SEQ ID No. 804	2188.2	Contig51	m	56047	56799	-	-	100%		-
SEQ ID No. 3290	767.3	Contig51	m	56793	58106	-	-	97%		-
SEQ ID No. 3291	769.2	Contig51	m	58073	59020	-	-	98%		-
SEQ ID No. 805	2191.4	Contig51	m	59014	59529	-	-	100%		-
SEQ ID No. 1620	3445.3	Contig51	p	59617	60657	+	-	99%		-
SEQ ID No. 1621	3446.1	Contig51	p	60510	61739	-	-	98%		-

SEQ ID No. 1622	3447.1	Contig51	p	61720	62166	-	-	97%		+
SEQ ID No. 2777	5200.2	Contig51	p	62234	64525	-	-	99%		-
SEQ ID No. 1503	3271.3	Contig51	m	64736	66802	-	-	99%		-
SEQ ID No. 1504	3272.1	Contig51	m	66845	67288	-	-	100%		-
SEQ ID No. 1505	3273.1	Contig51	p	67395	67838	-	-	100%		-
SEQ ID No. 1506	3275.1	Contig51	p	67835	69160	-	-	97%		+
SEQ ID No. 1507	3276.2	Contig51	m	69289	69438		-	93%		+
SEQ ID No. 1508	3277.2	Contig51	p	69443	70258	-	-	98%		-
SEQ ID No. 1509	3278.1	Contig51	m	70313	70750	-	-	98%		-
SEQ ID No. 1511	3280.1	Contig51	m	70844	71740	-	-	99%		-
SEQ ID No. 1512	3281.1	Contig51	p	71741	72040		-	97%		+
SEQ ID No. 3201	637.3	Contig51	m	71849	73939	-	-	99%		-
SEQ ID No. 3200	636.2	Contig51	m	74029	74943	-	-	100%		-
SEQ ID No. 3199	635.5	Contig51	m	75191	76447	-	-	100%		-
SEQ ID No. 3198	634.4	Contig51	p	75669	76205		-	99%		-
SEQ ID No. 1513	3282.3	Contig51	m	76582	77901	-	-	100%	+	-
SEQ ID No. 1514	3283.1	Contig51	m	77820	78542	-	-	100%		-
SEQ ID No. 354	1463.2	Contig51	m	78555	79679	-	-	99%	+	-
SEQ ID No. 353	1461.3	Contig51	m	79633	80634	-	-	99%		-
SEQ ID No. 907	2358.3	Contig51	m	80568	81986	-	-	99%		-
SEQ ID No. 512	1701.5	Contig51	m	81987	83204	+	-	99%		-
SEQ ID No. 2931	5527.2	Contig51	m	83168	83683	-	+			-
SEQ ID No. 511	1699.4	Contig51	m	83619	84545	-	-	97%		-
SEQ ID No. 1798	3694.2	Contig51	m	84535	85659	-	-	99%		-
SEQ ID No. 3231	683.2	Contig51	m	85735	87123	+	-	99%		-
SEQ ID No. 3232	684.5	Contig51	m	87313	88107	-	-	99%		-
SEQ ID No. 1799	3697.4	Contig51	m	88097	91603	-	-	98%		-
SEQ ID No. 1428	3150.1	Contig51	p	91763	92446		-	99%		-
SEQ ID No. 1426	3149.1	Contig51	m	92548	93645		-	99%		+
SEQ ID No. 1425	3148.1	Contig51	p	93871	94731	+	-	98%		-
SEQ ID No. 1424	3147.1	Contig51	m	94805	95329	-	-	100%		-
SEQ ID No. 3259	727.3	Contig51	m	95296	98532	+	-	99%		-
SEQ ID No. 750	2098.2	Contig51	m	98628	98909	-	-	95%		+
SEQ ID No. 749	2095.2	Contig51	p	99001	100194	-	-	95%		-
SEQ ID No. 1423	3146.2	Contig51	m	100234	101040	-	-	98%		-
SEQ ID No. 1422	3145.1	Contig51	p	100490	100798		-	98%		+

SEQ ID No. 3366	880.3	Contig51	m	100986	102164	-	-	99%	-
SEQ ID No. 3367	881.2	Contig51	m	102053	103141	+	-	98%	-
SEQ ID No. 3368	882.3	Contig51	p	103303	104787	-	-	98%	-
SEQ ID No. 824	2221.3	Contig51	p	104753	105202	-	-	97%	-
SEQ ID No. 2245	4373.3	Contig51	m	105349	105768	-	-	97%	-
SEQ ID No. 1033	2542.2	Contig51	m	105689	107065	-	-	98%	-
SEQ ID No. 2246	4374.4	Contig51	m	107058	108635	-	-	99%	-
SEQ ID No. 3054	5806.2	Contig51	p	108987	109253	-	-	100%	-
SEQ ID No. 2557	4818.3	Contig51	m	109314	110570	-	-	98%	+
SEQ ID No. 2556	4817.4	Contig51	m	110877	112283	-	-	98%	+
SEQ ID No. 2100	4167.4	Contig51	m	112666	116331	+	-	98%	-
SEQ ID No. 2099	4165.1	Contig51	m	116295	116864	+	-	97%	-
SEQ ID No. 2098	4164.1	Contig51	m	116858	117592	+	-	99%	-
SEQ ID No. 2097	4162.2	Contig51	p	117747	120260		-	97%	-
SEQ ID No. 3193	628.3	Contig51	p	120526	121188		-	98%	+
SEQ ID No. 3192	627.1	Contig51	p	121353	122504	-	-	93%	-
SEQ ID No. 820	2217.3	Contig51	m	122528	124399	-	-	99%	-
SEQ ID No. 384	1511.3	Contig51	m	124396	125073	-	-	99%	-
SEQ ID No. 383	1510.3	Contig51	m	125180	126028	-	-	99%	-
SEQ ID No. 381	1508.4	Contig51	p	126402	127535	-	-	98%	-
SEQ ID No. 1579	3385.1	Contig51	m	127494	128615	-	-	99%	-
SEQ ID No. 1580	3386.1	Contig51	m	128733	129023	-	-	100%	-
SEQ ID No. 641	1894.2	Contig51	m	129024	129398	-	-	100%	-
SEQ ID No. 640	1893.2	Contig51	p	129560	130237	-	-	100%	-
SEQ ID No. 842	2253.2	Contig51	p	130333	130923	-	-	98%	-
SEQ ID No. 841	2252.1	Contig51	p	130939	132033	-	-	99%	-
SEQ ID No. 1581	3388.1	Contig51	p	132395	133510	-	-	98%	-
SEQ ID No. 1582	3389.1	Contig51	p	133644	134612	-	-	99%	-
SEQ ID No. 1584	3390.1	Contig51	m	134639	134893		-	100%	-
SEQ ID No. 1585	3391.2	Contig51	m	134835	137117	-	-	99%	-
SEQ ID No. 2768	5180.1	Contig51	m	137087	137341	-	+		-
SEQ ID No. 1690	3544.2	Contig51	m	137403	138500	-	-	98%	-
SEQ ID No. 1689	3543.1	Contig51	m	138482	139159	-	-	96%	-
SEQ ID No. 1688	3542.1	Contig51	p	139288	140169	-	-	99%	-
SEQ ID No. 1687	3541.1	Contig51	p	140157	140939	-	-	98%	-
SEQ ID No. 1370	307.2	Contig51	p	140927	141469	-	-	97%	-

SEQ ID No. 1362	306.1	Contig51	m	141483	142922	-	-	99%		-
SEQ ID No. 1355	305.2	Contig51	m	142891	143622	-	-	99%		-
SEQ ID No. 527	1724.3	Contig51	p	143835	144755	-	-	98%		-
SEQ ID No. 528	1725.3	Contig51	m	144891	145793	-	-	100%		-
SEQ ID No. 2183	4294.3	Contig51	m	146041	147027	-	-	99%		-
SEQ ID No. 890	2337.3	Contig51	m	146964	147656	-	-	99%		-
SEQ ID No. 891	2339.4	Contig51	m	147601	148743	-	-	98%		-
SEQ ID No. 2184	4295.3	Contig51	p	148722	150056	-	-	99%		-
SEQ ID No. 2185	4296.1	Contig51	p	150005	151357	-	-	98%		-
SEQ ID No. 2186	4297.2	Contig51	p	151324	151899	-	-	97%		-
SEQ ID No. 2767	5178.1	Contig51	p	151974	152465	+	-	99%		-
SEQ ID No. 911	2362.3	Contig51	p	152437	153357	+	-	99%		-
SEQ ID No. 912	2364.1	Contig51	p	153306	154490	-	-	99%	+	-
SEQ ID No. 913	2365.2	Contig51	p	154521	156299	+	-	95%		-
SEQ ID No. 524	1720.3	Contig51	m	156341	157147	-	-	99%		-
SEQ ID No. 1684	3538.1	Contig51	m	157168	158121	-	-	98%		-
SEQ ID No. 239	1277.2	Contig51	p	158205	159371	-	-	99%		-
SEQ ID No. 238	1275.2	Contig51	p	159645	160532	-	-	97%		-
SEQ ID No. 237	1273.4	Contig51	p	160659	161378	-	-	90%		-
SEQ ID No. 1686	3540.2	Contig51	p	161362	163212	-	-	93%		-
SEQ ID No. 1715	3582.2	Contig51	p	163197	164072	-	-	89%		-
SEQ ID No. 1714	3581.1	Contig51	m	164333	165085	-	-	93%		-
SEQ ID No. 559	1770.2	Contig51	m	165180	168293	+	-	85%		+
SEQ ID No. 560	1772.1	Contig51	m	168284	168760	-	-	89%		-
SEQ ID No. 561	1773.1	Contig51	p	168849	169307	-	-	99%		-
SEQ ID No. 562	1774.2	Contig51	p	169304	170725	+	-	87%		-
SEQ ID No. 1713	3578.2	Contig51	m	170732	172456	-	-	88%		+
SEQ ID No. 1887	3846.2	Contig51	p	172718	175288	-	-	95%		-
SEQ ID No. 1150	273.1	Contig51	p	175515	176069	-	-	98%		+
SEQ ID No. 1139	271.1	Contig51	m	176141	177055	-	-	98%		-
SEQ ID No. 1132	270.1	Contig51	m	177065	177934	-	-	99%		-
SEQ ID No. 1126	269.1	Contig51	m	178010	179101	-	-	98%		-
SEQ ID No. 1120	268.1	Contig51	m	179040	179276	-	+			-
SEQ ID No. 1115	267.3	Contig51	m	179257	180225	+	-	98%		-
SEQ ID No. 2793	5232.1	Contig51	m	180191	181312	-	-	93%		-
SEQ ID No. 1403	3117.2	Contig51	m	181258	181983	-	+			-

SEQ ID No. 208	123.2	Contig51	m	181917	183347	+	-	99%		-
SEQ ID No. 216	124.1	Contig51	m	183344	183853	-	-	100%		-
SEQ ID No. 222	125.2	Contig51	m	183952	185502	-	-	98%		-
SEQ ID No. 707	2022.1	Contig51	p	185514	185687		-	98%		+
SEQ ID No. 706	2021.1	Contig51	m	185561	185980	-	-	99%		-
SEQ ID No. 705	2020.1	Contig51	m	185959	186597	-	-	99%		-
SEQ ID No. 703	2019.2	Contig51	m	186827	187576	+	-	99%		-
SEQ ID No. 3240	698.2	Contig51	m	187570	188787	-	-	99%		-
SEQ ID No. 3239	697.1	Contig51	m	188795	189424	+	-	98%		-
SEQ ID No. 3238	696.2	Contig51	m	189586	190029	-	-	100%		-
SEQ ID No. 2791	5229.1	Contig51	m	190023	190493	-	-	99%		-
SEQ ID No. 2578	4856.2	Contig51	m	190797	191123	-	-	96%		-
SEQ ID No. 2577	4855.2	Contig51	m	191116	191436	-	-	100%		-
SEQ ID No. 940	2407.3	Contig51	m	191412	193793	-	-	98%		-
SEQ ID No. 2336	4503.1	Contig51	m	194051	195133	-	-	99%	+	-
SEQ ID No. 2337	4505.3	Contig51	m	195204	195566	-	-	100%	+	-
SEQ ID No. 2790	5227.2	Contig51	m	195579	195845	-	+			-
SEQ ID No. 2789	5226.2	Contig51	m	195800	196339		-	99%		-
SEQ ID No. 2622	4923.2	Contig51	m	196361	198343	-	-	99%		-
SEQ ID No. 1669	3516.2	Contig51	p	198784	199077	-	-	100%		-
SEQ ID No. 1668	3515.2	Contig51	p	199397	199897	-	-	98%		-
SEQ ID No. 1667	3514.1	Contig51	p	199903	200154		+			+
SEQ ID No. 1666	3512.1	Contig51	m	200543	202108	-	-	97%		-
SEQ ID No. 1665	3511.2	Contig51	p	202374	203537	-	-	98%		-
SEQ ID No. 1663	3509.2	Contig51	p	203513	204547	-	-	98%		-
SEQ ID No. 1064	2590.3	Contig51	p	204548	205471	-	-	99%		-
SEQ ID No. 1065	2591.3	Contig51	m	205599	206756	-	-	98%		-
SEQ ID No. 1662	3508.2	Contig51	m	206726	207028	-	-	97%		-
SEQ ID No. 1661	3507.1	Contig51	m	207278	207652	-	-	98%		+
SEQ ID No. 1660	3506.1	Contig51	p	207724	208368	-	-	100%	+	-
SEQ ID No. 1659	3504.1	Contig51	m	208578	209129	-	-	99%	+	-
SEQ ID No. 1658	3503.2	Contig51	m	209099	210298	-	-	99%	+	-
SEQ ID No. 1657	3502.3	Contig51	m	210312	210839	+	-	97%		-
SEQ ID No. 3053	5804.1	Contig51	m	210800	211321	+	-	100%		-
SEQ ID No. 388	1517.3	Contig51	p	211384	212667	-	-	98%		-
SEQ ID No. 389	1519.2	Contig51	m	212748	213812	-	-	100%		-

SEQ ID No. 391	1521.2	Contig51	p	213997	214617	-	-	97%		-
SEQ ID No. 2276	4415.3	Contig51	m	214705	215997	-	-	100%		-
SEQ ID No. 2205	4325.3	Contig51	p	216364	217314	-	-	97%		+
SEQ ID No. 2206	4326.1	Contig51	p	217278	218063	-	-	96%		-
SEQ ID No. 2207	4327.1	Contig51	p	218064	219191	-	-	97%		-
SEQ ID No. 2208	4328.1	Contig51	p	219146	220381	+	-	98%		-
SEQ ID No. 2499	4740.2	Contig51	m	220517	221473	-	-	99%	+	-
SEQ ID No. 2497	4739.1	Contig51	m	221437	222102	-	-	99%		-
SEQ ID No. 2496	4737.1	Contig51	m	222089	222667	-	-	100%		-
SEQ ID No. 2495	4736.2	Contig51	m	222642	223916	-	-	98%		-
SEQ ID No. 2494	4734.2	Contig51	p	223928	224914	-	-	97%		-
SEQ ID No. 2887	5440.1	Contig51	m	224949	226127	-	-	97%		+
SEQ ID No. 3202	64.3	Contig51	m	226279	228249	-	-	97%		+
SEQ ID No. 3209	65.1	Contig51	m	228321	228539	-	+			+
SEQ ID No. 2951	5556.1	Contig52	p	224	484		+			+
SEQ ID No. 2950	5555.2	Contig52	p	358	1134		+			+
SEQ ID No. 3147	6052.1	Contig52	m	1145	1837	-	+			-
SEQ ID No. 2778	5201.2	Contig52	m	1785	2489	-	+			-
SEQ ID No. 2779	5202.2	Contig52	m	2453	2761	-	+			-
SEQ ID No. 2780	5204.2	Contig52	m	2762	3238		+			-
SEQ ID No. 2295	4444.2	Contig52	p	3519	4844	-	-	98%		-
SEQ ID No. 2294	4442.1	Contig52	m	4989	5945	-	-	98%		-
SEQ ID No. 2293	4441.1	Contig52	p	6110	6532	-	-	99%		-
SEQ ID No. 2292	4440.1	Contig52	p	6424	6957	-	-	98%		-
SEQ ID No. 621	1863.2	Contig52	p	6950	7645	-	-	99%		-
SEQ ID No. 620	1862.2	Contig52	p	7735	9072	-	-	96%		-
SEQ ID No. 3315	801.3	Contig52	p	9080	9829	-	-	94%		-
SEQ ID No. 3316	802.3	Contig52	p	9968	10765	-	-	99%		-
SEQ ID No. 3317	803.3	Contig52	p	10775	11470	+	-	99%		-
SEQ ID No. 2781	5206.1	Contig52	p	11494	12612	+	-	98%		-
SEQ ID No. 1421	3142.2	Contig52	p	12546	13499	-	-	94%		-
SEQ ID No. 401	1532.3	Contig52	p	13507	15498	-	-	97%		-
SEQ ID No. 402	1533.1	Contig52	p	15502	16737	-	-	97%		-
SEQ ID No. 220	1246.2	Contig52	p	16920	18671	-	-	80%		-
SEQ ID No. 219	1245.1	Contig52	p	18783	19034		-	87%		+
SEQ ID No. 218	1244.2	Contig52	p	19306	20388		-	89%		+

SEQ ID No. 329	1417.1	Contig52	p	20421	20891	-	-	100%		-
SEQ ID No. 330	1418.3	Contig52	m	20994	22823	+	-	90%		-
SEQ ID No. 3146	6049.1	Contig52	m	22855	23631	-	+			-
SEQ ID No. 1285	2954.3	Contig52	m	23498	24508	-	-	99%		-
SEQ ID No. 3131	601.4	Contig52	p	24684	25916	-	+			-
SEQ ID No. 3127	600.3	Contig52	p	25859	26539	-	+			-
SEQ ID No. 3118	598.2	Contig52	p	26619	26939	+	-	100%		-
SEQ ID No. 3115	597.2	Contig52	p	26966	27346	+	-	98%		-
SEQ ID No. 3114	596.3	Contig52	m	27403	29127	-	-	99%	+	-
SEQ ID No. 1286	2955.2	Contig52	p	29162	29581	-	-	99%		+
SEQ ID No. 1287	2958.2	Contig52	p	29575	30342	-	-	98%		-
SEQ ID No. 1288	2959.1	Contig52	p	30346	31104	+	-	97%		-
SEQ ID No. 1290	2960.1	Contig52	p	31112	32215	-	-	98%		-
SEQ ID No. 1291	2961.1	Contig52	m	32318	33571	-	-	98%		-
SEQ ID No. 1292	2962.1	Contig52	p	33682	34458	-	-	98%		-
SEQ ID No. 1006	2495.2	Contig52	p	34434	34982	-	-	98%		-
SEQ ID No. 1005	2493.2	Contig52	p	34913	35578	-	-	98%		-
SEQ ID No. 1004	2492.2	Contig52	p	35616	35948	+	-	95%		+
SEQ ID No. 3409	938.3	Contig52	m	35989	37614	-	-	90%		-
SEQ ID No. 3408	937.3	Contig52	m	37875	39305	+	-	99%		-
SEQ ID No. 437	1589.3	Contig52	m	39287	39994	-	-	99%		-
SEQ ID No. 439	1590.1	Contig52	m	40025	40582	+	-	98%		-
SEQ ID No. 440	1591.4	Contig52	m	40813	42264		-	98%		-
SEQ ID No. 1293	2963.2	Contig52	m	42366	43133	-	-	98%		-
SEQ ID No. 1294	2964.1	Contig52	m	43134	43991	-	-	98%		-
SEQ ID No. 1295	2965.1	Contig52	p	44163	44477	-	-	100%		+
SEQ ID No. 1296	2966.1	Contig52	m	44480	45376	-	-	97%		+
SEQ ID No. 1297	2967.1	Contig52	m	45421	45705	-	-	100%		+
SEQ ID No. 1298	2968.3	Contig52	p	45890	48583	-	-	76%		-
SEQ ID No. 2892	5455.1	Contig52	m	48635	49105	-	-	98%		-
SEQ ID No. 787	2156.3	Contig52	m	49102	50409	-	-	98%		-
SEQ ID No. 538	1740.3	Contig52	p	50644	51432	-	-	100%		-
SEQ ID No. 536	1739.2	Contig52	p	51433	52056	-	-	99%		-
SEQ ID No. 535	1738.3	Contig52	p	52040	53257	-	-	99%	+	-
SEQ ID No. 408	1547.2	Contig52	p	53250	54101	-	-	99%		-
SEQ ID No. 407	1545.3	Contig52	p	54083	55753	-	-	99%	+	-

SEQ ID No. 1531	3311.1	Contig52	p	55754	57133	-	-	98%	+	-
SEQ ID No. 831	2235.2	Contig52	p	57282	59225		-	97%		-
SEQ ID No. 1532	3312.1	Contig52	p	59352	59669	-	-	98%		-
SEQ ID No. 2937	5534.2	Contig52	m	59770	61257	-	-	99%		-
SEQ ID No. 2121	4197.2	Contig52	m	61258	63633		-	99%		-
SEQ ID No. 2938	5535.1	Contig52	p	62267	62506		-	97%		+
SEQ ID No. 345	1448.2	Contig52	p	63680	64834	-	-	96%		-
SEQ ID No. 346	1449.3	Contig52	m	64855	67023	+	-	98%		-
SEQ ID No. 2120	4195.3	Contig52	m	67104	68273	+	-	99%		-
SEQ ID No. 2734	5100.2	Contig52	m	68261	69484	-	-	96%		-
SEQ ID No. 1578	3378.2	Contig52	m	69468	71318	-	-	98%		-
SEQ ID No. 1577	3376.1	Contig52	m	71343	72434	-	-	96%		-
SEQ ID No. 1773	366.2	Contig52	p	72814	73131	-	-	91%		+
SEQ ID No. 1788	368.2	Contig52	m	73211	75703		-	93%		-
SEQ ID No. 1576	3374.1	Contig52	m	75741	76907	-	-	94%		-
SEQ ID No. 3441	982.3	Contig52	m	77475	78893	+	-	98%	+	-
SEQ ID No. 3442	983.1	Contig52	m	78868	79611	-	-	98%		-
SEQ ID No. 3443	985.2	Contig52	m	79583	80569	-	-	99%		-
SEQ ID No. 1575	3371.2	Contig52	m	80548	81924	-	-	100%		-
SEQ ID No. 2553	4813.2	Contig52	m	81928	83229	-	-	99%		-
SEQ ID No. 1573	3368.2	Contig52	p	82091	82516		-	97%		+
SEQ ID No. 2554	4814.2	Contig52	m	83467	84225	-	-	99%		-
SEQ ID No. 2555	4815.2	Contig52	m	84536	84952	-	-	99%		-
SEQ ID No. 255	130.4	Contig52	m	84956	86611	-	-	98%		-
SEQ ID No. 249	129.3	Contig52	m	86644	86940	-	-	96%		+
SEQ ID No. 227	126.2	Contig52	m	87008	88483	-	-	98%		-
SEQ ID No. 2433	465.2	Contig52	p	88842	89729	-	-	99%		-
SEQ ID No. 2425	464.2	Contig52	p	89686	90996	-	-	95%		-
SEQ ID No. 2420	463.2	Contig52	m	90941	91528		-	95%		+
SEQ ID No. 2412	462.2	Contig52	p	90969	91781	+	-	94%		-
SEQ ID No. 2400	460.2	Contig52	p	91747	92295	-	-	98%		-
SEQ ID No. 2393	459.3	Contig52	m	92270	92908	-	-	99%		-
SEQ ID No. 2386	458.3	Contig52	m	92893	93930	-	-	98%		-
SEQ ID No. 1279	2943.1	Contig52	m	93920	94501	+	-	100%		-
SEQ ID No. 1278	2942.1	Contig52	m	94512	96995	-	-	99%		-
SEQ ID No. 1277	2941.1	Contig52	m	97049	98665		-	97%		-

SEQ ID No. 3108	5944.1	Contig52	p	98412	98600		-	96%		+
SEQ ID No. 198	1215.3	Contig52	p	98687	99826	-	-	100%		-
SEQ ID No. 200	1220.3	Contig52	p	99827	101356	-	-	98%		-
SEQ ID No. 751	2103.2	Contig52	p	101440	103845	+	-	99%		-
SEQ ID No. 470	1638.2	Contig52	p	103846	105045	-	-	99%	+	-
SEQ ID No. 471	1639.4	Contig52	m	105158	106336	-	-	91%		+
SEQ ID No. 468	1635.4	Contig52	m	106495	109404	-	-	92%		+
SEQ ID No. 2909	5481.1	Contig52	p	109609	110751	+	-	97%		-
SEQ ID No. 2445	4665.2	Contig52	m	110890	113277	-	-	98%		-
SEQ ID No. 2178	4288.2	Contig52	m	113583	114581	+	-	97%		-
SEQ ID No. 2179	4289.2	Contig52	m	114538	115254	+	-	99%		-
SEQ ID No. 2181	4291.1	Contig52	m	115152	115556	+	-	98%		-
SEQ ID No. 3293	770.3	Contig52	m	115526	116104	-	-	97%		-
SEQ ID No. 3294	771.4	Contig52	m	115998	116444	-	-	100%		-
SEQ ID No. 3295	772.4	Contig52	m	116524	117759	-	-	99%	+	-
SEQ ID No. 3296	774.3	Contig52	p	117790	119229	-	-	100%	+	-
SEQ ID No. 3297	775.2	Contig52	p	119230	119802	+	-	99%		-
SEQ ID No. 2182	4293.2	Contig52	p	119803	120579	-	-	99%		-
SEQ ID No. 2910	5482.2	Contig52	p	120600	121508	-	-	98%		-
SEQ ID No. 1920	3909.3	Contig52	p	121603	123297	-	-	97%		+
SEQ ID No. 1921	3911.2	Contig52	p	123484	125370	-	-	99%		-
SEQ ID No. 925	2382.2	Contig52	m	125482	125799	-	-	100%		-
SEQ ID No. 924	2381.2	Contig52	m	125894	127072	-	-	96%		-
SEQ ID No. 922	2377.2	Contig52	m	127035	128015	-	-	99%		-
SEQ ID No. 921	2375.3	Contig52	p	128218	128793	-	-	98%		-
SEQ ID No. 1922	3913.2	Contig52	m	128905	130032	+	-	100%		-
SEQ ID No. 1923	3914.3	Contig52	m	130020	131921	+	-	99%		-
SEQ ID No. 832	2239.3	Contig52	p	132081	133076	-	+			-
SEQ ID No. 833	2240.2	Contig52	m	133339	133824	-	-	98%		-
SEQ ID No. 834	2241.2	Contig52	m	133818	134192	-	-	99%		-
SEQ ID No. 2253	4384.2	Contig52	m	134261	135727	-	-	99%		-
SEQ ID No. 3145	6044.1	Contig52	p	135288	135446		-	96%		+
SEQ ID No. 2254	4385.1	Contig52	p	135914	137590	-	-	99%	+	-
SEQ ID No. 2255	4386.2	Contig52	p	137514	139124	-	-	98%		-
SEQ ID No. 2585	4867.2	Contig52	p	139087	140526	-	-	92%		-
SEQ ID No. 2586	4868.2	Contig52	p	140406	141347	-	-	98%		-

SEQ ID No. 2587	4869.3	Contig52	p	141335	141775	-	-	100%		-
SEQ ID No. 2945	5549.2	Contig52	p	141660	142481	-	-	99%		-
SEQ ID No. 2944	5548.1	Contig52	m	142688	143068	-	-	100%		+
SEQ ID No. 2943	5546.1	Contig52	p	143354	143830	-	-	100%		-
SEQ ID No. 2590	4874.2	Contig52	p	143907	145352	-	-	77%		+
SEQ ID No. 2591	4875.1	Contig52	m	145418	145843	-	-	96%		-
SEQ ID No. 2592	4877.2	Contig52	m	145945	147057	-	-	98%		-
SEQ ID No. 2419	4629.2	Contig52	p	147294	147728	-	-	98%		+
SEQ ID No. 2418	4628.2	Contig52	p	147706	148005	-	-	100%		-
SEQ ID No. 2417	4627.2	Contig52	p	147984	149039	-	-	100%		-
SEQ ID No. 2416	4626.1	Contig52	p	149012	149989	-	-	99%		-
SEQ ID No. 2415	4625.2	Contig52	p	149995	150963	-	-	98%		-
SEQ ID No. 2545	4803.2	Contig52	p	150975	151721	-	-	99%		-
SEQ ID No. 2546	4804.1	Contig52	p	151758	152054	-	-	100%		-
SEQ ID No. 2547	4805.1	Contig52	p	152038	153312	+	-	99%		-
SEQ ID No. 2548	4806.4	Contig52	p	153313	154311	+	-	97%		-
SEQ ID No. 3144	6042.1	Contig52	p	154287	154949	-	-	97%		-
SEQ ID No. 2641	4958.3	Contig52	p	154943	155851	-	-	96%		-
SEQ ID No. 2640	4957.2	Contig52	p	155962	156318	-	-	98%		-
SEQ ID No. 282	1346.3	Contig52	p	156342	157166	-	-	98%		-
SEQ ID No. 281	1345.3	Contig52	p	157275	158378	-	-	97%		-
SEQ ID No. 1670	3517.3	Contig52	p	158419	159726	-	-	99%		-
SEQ ID No. 347	1450.2	Contig52	m	160244	161539	-	-	98%		-
SEQ ID No. 1671	3519.1	Contig52	p	161680	162921	-	-	98%		-
SEQ ID No. 1672	3520.1	Contig52	m	163009	164394	+	-	97%		-
SEQ ID No. 1673	3521.2	Contig52	p	164484	165635	-	-	97%		-
SEQ ID No. 1674	3522.1	Contig52	m	165688	166524	-	-	99%		-
SEQ ID No. 1675	3524.1	Contig52	p	166701	167300	-	-	98%		-
SEQ ID No. 1676	3525.1	Contig52	p	167624	168361	-	-	99%		-
SEQ ID No. 1677	3527.1	Contig52	p	168334	168732	-	-	100%		-
SEQ ID No. 2139	4228.2	Contig52	m	168692	170233	-	-	99%		-
SEQ ID No. 2138	4227.1	Contig52	p	170259	171749	-	-	98%		-
SEQ ID No. 96	1059.2	Contig52	m	171843	173126	-	-	99%	+	-
SEQ ID No. 97	1060.2	Contig52	p	173322	174221	-	-	94%		-
SEQ ID No. 2137	4226.2	Contig52	p	174266	176896	-	-	98%		-
SEQ ID No. 932	2395.3	Contig52	p	176847	177977	-	-	94%		-

SEQ ID No. 454	1611.5	Contig52	p	177956	179017	-	-	98%	-
SEQ ID No. 3143	6040.1	Contig52	p	178884	179285	-	+		-
SEQ ID No. 3081	5876.2	Contig52	p	179252	179977	-	-	99%	-
SEQ ID No. 3141	6039.1	Contig52	p	179971	181719	-	-	100%	-
SEQ ID No. 2966	5588.2	Contig52	p	181821	182123	+	-	100%	-
SEQ ID No. 2325	4489.3	Contig52	p	182170	183339	-	-	100%	-
SEQ ID No. 2324	4488.1	Contig52	p	183570	184718	-	-	100%	-
SEQ ID No. 2323	4487.1	Contig52	p	184719	185408	-	-	97%	-
SEQ ID No. 578	1795.3	Contig52	p	185529	188249	-	-	98%	-
SEQ ID No. 2542	4800.2	Contig52	p	188179	188973	-	-	98%	-
SEQ ID No. 2543	4801.1	Contig52	p	188964	189461		-	93%	-
SEQ ID No. 2544	4802.2	Contig52	p	189554	190114	-	-	94%	-
SEQ ID No. 2965	5587.3	Contig52	p	190081	190959	-	-	98%	-
SEQ ID No. 2735	5103.4	Contig52	m	191093	191497	-	-	100%	-
SEQ ID No. 2736	5104.3	Contig52	m	191571	193376		-	97%	-
SEQ ID No. 577	1794.3	Contig52	m	193397	194182	-	-	97%	-
SEQ ID No. 576	1793.3	Contig52	m	194158	195009	-	-	99%	-
SEQ ID No. 2894	5457.1	Contig52	m	195010	195438	-	-	99%	-
SEQ ID No. 2670	5003.2	Contig52	m	195444	196262		-	97%	+
SEQ ID No. 2671	5005.3	Contig52	m	196424	197824	+	-	99%	-
SEQ ID No. 2893	5456.1	Contig52	m	197784	198467	-	-	99%	-
SEQ ID No. 1812	3712.2	Contig52	m	198472	199365	-	-	98%	-
SEQ ID No. 1811	3711.1	Contig52	m	199322	199801	-	-	100%	-
SEQ ID No. 1810	3710.1	Contig52	m	199795	200760	-	-	98%	-
SEQ ID No. 1808	3709.1	Contig52	p	201015	201863	-	-	96%	-
SEQ ID No. 1807	3708.1	Contig52	p	201767	202480	-	-	99%	-
SEQ ID No. 1806	3706.2	Contig52	p	202699	203940	-	-	99%	-
SEQ ID No. 1805	3705.1	Contig52	p	203934	204731	-	-	98%	-
SEQ ID No. 1804	3703.1	Contig52	p	204682	205314	-	-	98%	-
SEQ ID No. 1803	3702.1	Contig52	p	205289	206050	-	-	99%	-
SEQ ID No. 1802	3701.1	Contig52	p	206268	207071		-	97%	-
SEQ ID No. 1800	3699.3	Contig52	p	207457	210069		-	96%	+
SEQ ID No. 2149	4242.2	Contig52	p	210147	211148	-	-	95%	-
SEQ ID No. 2150	4243.2	Contig52	m	211272	211772	-	-	95%	-
SEQ ID No. 2151	4244.1	Contig52	p	211897	212304	+	+		+
SEQ ID No. 2152	4246.2	Contig52	p	212331	213194	-	-	96%	-

SEQ ID No. 266	1323.2	Contig52	p	213309	213839	-	-	98%		+
SEQ ID No. 265	1322.2	Contig52	m	213961	215397	+	-	97%		-
SEQ ID No. 395	1525.3	Contig52	m	215375	216646		-	98%		-
SEQ ID No. 396	1528.4	Contig52	p	216734	218086	-	-	98%		-
SEQ ID No. 400	1531.3	Contig52	p	217993	220296	-	-	99%		-
SEQ ID No. 399	1530.2	Contig52	m	220337	220723	-	-	98%		-
SEQ ID No. 397	1529.2	Contig52	p	220737	221954	-	-	99%		-
SEQ ID No. 1540	3327.1	Contig52	p	221944	222768	-	-	96%		+
SEQ ID No. 1541	3328.1	Contig52	p	222755	224497	-	-	98%		-
SEQ ID No. 1543	3331.1	Contig52	p	224570	225568	-	-	96%		-
SEQ ID No. 433	1584.3	Contig52	p	225621	228359	-	-	99%	+	+
SEQ ID No. 1544	3333.1	Contig52	p	228335	228760	-	-	99%		-
SEQ ID No. 1545	3334.2	Contig52	p	228944	229807		-	97%		+
SEQ ID No. 2678	5021.3	Contig52	m	229852	230601		-	100%		-
SEQ ID No. 2676	5019.1	Contig52	m	230836	231474	-	-	99%		-
SEQ ID No. 2675	5018.2	Contig52	m	231413	231898		-	98%		+
SEQ ID No. 2464	4696.3	Contig52	m	231883	233508		-	99%		-
SEQ ID No. 2463	4695.1	Contig52	p	233502	233720		+			+
SEQ ID No. 2462	4694.1	Contig52	p	233593	235266	-	-	98%		-
SEQ ID No. 2461	4692.5	Contig52	m	235587	236966	-	-	98%		-
SEQ ID No. 444	1598.6	Contig52	p	236937	237929	-	-	100%		-
SEQ ID No. 443	1595.3	Contig52	p	237930	239072	-	-	98%		-
SEQ ID No. 442	1594.2	Contig52	p	239053	239775	-	-	97%		-
SEQ ID No. 989	2472.2	Contig52	p	239763	240410	-	-	96%		-
SEQ ID No. 990	2473.1	Contig52	p	240545	240859	-	-	100%		-
SEQ ID No. 991	2474.1	Contig52	p	240853	241485	-	-	99%		-
SEQ ID No. 2609	4901.3	Contig52	p	241486	243285	-	-	99%		-
SEQ ID No. 2608	4900.3	Contig52	p	243257	246211		-	98%		+
SEQ ID No. 2247	4376.3	Contig52	p	246162	246848	-	-	97%		-
SEQ ID No. 2248	4377.4	Contig52	m	246921	247985	-	-	96%		+
SEQ ID No. 3140	6038.1	Contig52	p	247180	247419		-	96%		+
SEQ ID No. 3433	971.2	Contig52	m	248160	249020	-	-	98%		+
SEQ ID No. 3434	972.3	Contig52	p	249094	250731	-	-	96%		-
SEQ ID No. 2604	4891.2	Contig52	m	250785	251615	-	-	99%		-
SEQ ID No. 2602	4889.1	Contig52	m	251793	252218	-	-	100%		-
SEQ ID No. 2601	4888.1	Contig52	m	252280	252804	-	-	96%		-

SEQ ID No. 2600	4887.3	Contig52	p	252888	253451	-	-	97%	-
SEQ ID No. 2716	5075.4	Contig52	p	253679	256276	-	-	94%	-
SEQ ID No. 3139	6037.1	Contig52	p	255546	255887		-	81%	-
SEQ ID No. 2715	5072.2	Contig52	m	255557	256015		-	82%	-
SEQ ID No. 2714	5071.2	Contig52	p	256486	257511	-	-	88%	+
SEQ ID No. 2296	4452.2	Contig52	p	257614	259002		-	94%	-
SEQ ID No. 2297	4453.2	Contig52	p	259133	260563	-	-	78%	+
SEQ ID No. 2298	4454.4	Contig52	m	260752	261759	-	-	95%	-
SEQ ID No. 3138	6036.1	Contig52	p	262003	262335	-	+		+
SEQ ID No. 2963	5584.2	Contig52	m	262416	263513	-	+		+
SEQ ID No. 2662	4993.2	Contig52	m	263713	264231		-	99%	-
SEQ ID No. 2661	4992.2	Contig52	p	264173	264472	-	-	98%	+
SEQ ID No. 2660	4990.3	Contig52	p	264835	266655		-	94%	+
SEQ ID No. 984	2466.3	Contig52	p	266848	269556	-	-	95%	-
SEQ ID No. 2389	4584.3	Contig52	m	269691	270662	-	-	98%	+
SEQ ID No. 2953	5559.3	Contig52	m	270701	271255	-	-	97%	+
SEQ ID No. 2954	5560.3	Contig52	m	271292	271954	-	-	98%	-
SEQ ID No. 2955	5563.1	Contig52	m	272495	272959	-	-	100%	-
SEQ ID No. 2956	5564.1	Contig52	m	273293	274750	-	-	99%	-
SEQ ID No. 3387	908.4	Contig52	m	274806	276458	-	-	97%	-
SEQ ID No. 2237	4364.2	Contig52	p	274900	275220		-	97%	-
SEQ ID No. 3386	907.4	Contig52	m	276498	279179	-	-	99%	-
SEQ ID No. 2388	4581.2	Contig52	m	279189	279581		-	99%	+
SEQ ID No. 2387	4580.1	Contig52	m	279617	280120		-	90%	-
SEQ ID No. 2385	4578.1	Contig52	m	280373	281584	-	-	99%	-
SEQ ID No. 2384	4577.2	Contig52	m	281864	282706		-	98%	-
SEQ ID No. 2258	4389.2	Contig52	p	282782	284104	+	-	99%	-
SEQ ID No. 2259	4391.1	Contig52	p	284105	284953	-	-	97%	-
SEQ ID No. 2260	4392.1	Contig52	p	284954	285220	-	+		-
SEQ ID No. 2261	4393.1	Contig52	p	285109	285819	-	-	96%	-
SEQ ID No. 2262	4394.4	Contig52	p	285823	287865	-	-	98%	-
SEQ ID No. 2699	5052.4	Contig52	p	287980	288597	-	-	96%	+
SEQ ID No. 2698	5050.2	Contig52	p	288634	289383	-	-	98%	-
SEQ ID No. 1883	3838.3	Contig52	p	289389	291629	-	-	99%	-
SEQ ID No. 1885	3840.1	Contig52	p	291630	292769	-	-	99%	-
SEQ ID No. 1886	3841.2	Contig52	p	292848	295433	-	-	85%	-

SEQ ID No. 3341	843.2	Contig52	m	295435	296004	-	-	99%	-
SEQ ID No. 3342	844.2	Contig52	m	296047	296436	-	-	99%	-
SEQ ID No. 3343	845.4	Contig52	m	296556	297314		-	98%	-
SEQ ID No. 1002	249.4	Contig52	p	297549	299282	-	-	98%	-
SEQ ID No. 3071	5848.1	Contig52	m	298199	298495		-	96%	+
SEQ ID No. 1008	250.1	Contig52	p	299279	300508	-	-	99%	-
SEQ ID No. 1014	251.2	Contig52	p	300542	301417	-	-	99%	+
SEQ ID No. 1019	252.2	Contig52	m	301431	302669	-	-	98%	-
SEQ ID No. 1912	3898.2	Contig52	m	302744	303340	-	-	98%	+
SEQ ID No. 1913	3899.2	Contig52	m	303353	303835		-	98%	+
SEQ ID No. 1916	3901.2	Contig52	p	303849	304757	-	-	99%	-
SEQ ID No. 1917	3902.3	Contig52	m	304877	306370	-	-	98%	-
SEQ ID No. 1893	3857.3	Contig52	m	306327	307457	+	-	98%	-
SEQ ID No. 1892	3855.1	Contig52	m	307661	308464	-	-	98%	-
SEQ ID No. 315	1398.2	Contig52	m	308532	310538	-	-	97%	-
SEQ ID No. 316	1399.3	Contig52	m	310525	312153	-	-	99%	-
SEQ ID No. 1891	3853.2	Contig52	m	312450	313760	-	-	98%	-
SEQ ID No. 1890	3851.1	Contig52	p	314121	314300	-	+		-
SEQ ID No. 1889	3850.1	Contig52	m	314297	314740	-	-	97%	-
SEQ ID No. 1888	3849.1	Contig52	m	314722	315453	-	-	99%	-
SEQ ID No. 2582	4863.4	Contig52	p	315679	317031	-	-	99%	-
SEQ ID No. 2583	4864.2	Contig52	p	317013	317423	-	+		-
SEQ ID No. 2891	5453.2	Contig53	m	2	1057	+	-	99%	-
SEQ ID No. 1486	3244.2	Contig53	p	1210	1977	-	-	99%	-
SEQ ID No. 3187	618.3	Contig53	p	2045	2461	-	-	100%	-
SEQ ID No. 3188	619.1	Contig53	p	2371	3357	-	-	97%	-
SEQ ID No. 3189	621.5	Contig53	p	3335	4654	-	-	99%	-
SEQ ID No. 3233	687.5	Contig53	p	4655	5932	+	-	98%	-
SEQ ID No. 3234	688.1	Contig53	p	5916	6770	-	-	99%	-
SEQ ID No. 1485	3243.1	Contig53	m	6880	7188	-	-	99%	+
SEQ ID No. 1484	3242.3	Contig53	m	7361	10084	-	-	98%	+
SEQ ID No. 2989	5633.2	Contig53	m	10360	11202	-	-	98%	-
SEQ ID No. 2990	5634.2	Contig53	m	11203	11337	+	+		-
SEQ ID No. 2991	5636.1	Contig53	m	11490	11912		-	99%	+
SEQ ID No. 2992	5637.1	Contig53	p	12003	12689	-	-	99%	-
SEQ ID No. 2516	4766.2	Contig53	p	12665	13120	-	-	100%	-

SEQ ID No. 2517	4767.2	Contig53	p	13062	13379	-	-	100%		-
SEQ ID No. 2518	4768.1	Contig53	m	13467	13820		-	99%		-
SEQ ID No. 2519	4769.1	Contig53	p	14000	14419	-	-	100%		-
SEQ ID No. 2520	4770.1	Contig53	p	14512	14700	-	+			-
SEQ ID No. 2521	4771.2	Contig53	p	14837	15913	-	-	96%		-
SEQ ID No. 2993	5638.2	Contig53	m	16224	16634	+	-	100%		+
SEQ ID No. 2854	5377.2	Contig53	p	16834	17466	-	-	95%		+
SEQ ID No. 2853	5376.2	Contig53	m	17524	21840	-	-	99%		+
SEQ ID No. 1375	3077.1	Contig53	m	21958	22719	-	-	97%		-
SEQ ID No. 1376	3078.2	Contig53	m	22682	23938	-	-	99%		-
SEQ ID No. 1377	3080.2	Contig53	m	23895	25349	-	-	100%		-
SEQ ID No. 1378	3081.2	Contig53	m	25561	26097	-	-	98%		+
SEQ ID No. 1379	3082.1	Contig53	m	26262	26786	-	-	98%		-
SEQ ID No. 1380	3083.1	Contig53	m	26732	27307	+	-	97%		-
SEQ ID No. 840	2251.3	Contig53	m	27542	28039	-	-	92%		+
SEQ ID No. 1381	3085.2	Contig53	p	28077	30950	-	-	98%	+	-
SEQ ID No. 1382	3086.1	Contig53	p	31106	31714	+	-	100%		-
SEQ ID No. 1383	3087.1	Contig53	p	31707	32744	-	-	99%		-
SEQ ID No. 1384	3088.1	Contig53	m	32771	33550	-	-	100%		-
SEQ ID No. 2594	488.2	Contig53	m	33534	34451	-	-	99%		-
SEQ ID No. 2603	489.1	Contig53	m	34530	34715	-	+			+
SEQ ID No. 2613	491.1	Contig53	m	34679	35737	-	-	97%		+
SEQ ID No. 2620	492.4	Contig53	m	35824	36363	-	-	98%		-
SEQ ID No. 2852	5373.1	Contig53	m	36255	36521		+			+
SEQ ID No. 2625	493.4	Contig53	m	36559	37296	-	-	96%		-
SEQ ID No. 1386	3091.1	Contig53	m	37453	37764	-	-	100%		+
SEQ ID No. 3075	5867.1	Contig53	p	37577	37750		-	94%		+
SEQ ID No. 515	1705.2	Contig53	p	37874	38746	-	-	98%		+
SEQ ID No. 516	1706.2	Contig53	m	38847	39317	-	-	98%		-
SEQ ID No. 517	1707.2	Contig53	m	39318	39686	-	-	100%		-
SEQ ID No. 1387	3093.1	Contig53	m	39710	40480	-	-	98%		-
SEQ ID No. 1388	3095.1	Contig53	m	40462	41001	-	-	98%		-
SEQ ID No. 1389	3096.1	Contig53	m	40977	41246	-	-	100%		-
SEQ ID No. 1390	3097.3	Contig53	m	41324	42712	-	-	99%		-
SEQ ID No. 1947	3960.2	Contig53	m	42957	43676	-	-	93%		+
SEQ ID No. 1945	3959.2	Contig53	p	43842	45401		-	97%		-

SEQ ID No. 3270	742.3	Contig53	p	45729	47141	+	-	99%		-
SEQ ID No. 3269	741.2	Contig53	p	47095	47721	-	-	97%		+
SEQ ID No. 3268	740.2	Contig53	p	47919	48266	-	-	100%		-
SEQ ID No. 819	2213.2	Contig53	m	48480	48932	+	-	100%		+
SEQ ID No. 1944	3958.1	Contig53	m	48942	49568	+	-	97%		-
SEQ ID No. 1943	3957.1	Contig53	m	49907	50113	+	+			-
SEQ ID No. 1942	3956.2	Contig53	p	50175	50906		-	98%		-
SEQ ID No. 2929	5524.1	Contig53	p	50884	51435	-	-	99%		-
SEQ ID No. 2928	5523.2	Contig53	p	51398	52447	+	-	97%		-
SEQ ID No. 935	2400.2	Contig53	p	52389	53321	-	-	99%		-
SEQ ID No. 936	2401.2	Contig53	p	53322	53972	+	-	98%		-
SEQ ID No. 361	1476.2	Contig53	p	54335	55246	-	-	100%		+
SEQ ID No. 360	1475.2	Contig53	m	55250	55540	-	-	100%		+
SEQ ID No. 690	1989.2	Contig53	p	55808	57301	-	-	98%		-
SEQ ID No. 692	1990.1	Contig53	p	57129	58001	-	-	97%		-
SEQ ID No. 942	241.2	Contig53	p	57980	59827	-	-	98%		-
SEQ ID No. 934	240.1	Contig53	p	59751	60809	-	-	99%		-
SEQ ID No. 929	239.1	Contig53	p	60784	61458	-	-	97%		-
SEQ ID No. 923	238.1	Contig53	p	61569	62993	-	-	98%	+	-
SEQ ID No. 908	236.1	Contig53	m	62888	63082		-	98%		+
SEQ ID No. 898	235.2	Contig53	m	63083	64465	-	-	98%		-
SEQ ID No. 1299	2969.1	Contig53	m	64558	64863	-	-	98%		-
SEQ ID No. 1301	2970.2	Contig53	m	64881	65435	-	-	100%		-
SEQ ID No. 3392	916.3	Contig53	m	65548	66546	-	-	98%		-
SEQ ID No. 3393	917.3	Contig53	m	66654	66896	-	-	100%		-
SEQ ID No. 3394	918.3	Contig53	p	66802	66963		-	100%		+
SEQ ID No. 3395	919.1	Contig53	p	67100	67537	-	-	92%		-
SEQ ID No. 3396	920.2	Contig53	m	67640	68068	-	-	97%		+
SEQ ID No. 2437	4655.3	Contig53	p	68083	69438	+	-	99%		-
SEQ ID No. 283	1353.3	Contig53	p	69533	70615	-	-	99%		+
SEQ ID No. 284	1354.2	Contig53	p	70542	71525	-	-	99%		+
SEQ ID No. 2436	4653.2	Contig53	p	71489	72391	-	-	98%		-
SEQ ID No. 2435	4652.2	Contig53	p	72395	72793	-	-	97%		-
SEQ ID No. 2287	4432.2	Contig53	m	72795	73247	-	-	97%		+
SEQ ID No. 2288	4433.1	Contig53	p	73488	74099	-	-	98%		-
SEQ ID No. 2289	4434.1	Contig53	m	74253	75062	-	-	92%		+

SEQ ID No. 2290	4435.4	Contig53	p	75252	77324	-	-	94%		+
SEQ ID No. 2339	451.3	Contig53	m	78146	79099	-	-	97%		+
SEQ ID No. 2334	450.1	Contig53	p	79261	79500	-	+			-
SEQ ID No. 2316	448.1	Contig53	p	79662	79931	-	-	100%		-
SEQ ID No. 2307	447.2	Contig53	p	79932	80276	-	-	100%		-
SEQ ID No. 2291	444.2	Contig53	p	80346	80738	-	-	96%		+
SEQ ID No. 2285	443.3	Contig53	p	80802	81404	-	-	99%		-
SEQ ID No. 1703	3562.3	Contig53	p	81579	82721	-	-	99%		-
SEQ ID No. 1702	3561.2	Contig53	p	82655	85069	-	-	99%		-
SEQ ID No. 3362	874.2	Contig53	p	85470	86042	+	-	99%		-
SEQ ID No. 3361	873.2	Contig53	p	86027	86338	+	-	96%		-
SEQ ID No. 3360	872.2	Contig53	p	86339	86992	-	-	100%		-
SEQ ID No. 3359	870.3	Contig53	p	86938	88077	+	-	98%		-
SEQ ID No. 3358	869.2	Contig53	m	87390	87602		-	94%		+
SEQ ID No. 3222	668.4	Contig53	p	88067	91228	+	-	97%		-
SEQ ID No. 3221	667.3	Contig53	m	89749	90402		-	95%		-
SEQ ID No. 1536	3322.2	Contig53	p	91234	92052	-	-	98%		-
SEQ ID No. 1537	3323.1	Contig53	p	92027	92644	-	-	100%		-
SEQ ID No. 1538	3324.1	Contig53	p	92655	93068	+	-	99%		-
SEQ ID No. 2217	434.3	Contig53	p	93248	93892	-	-	99%		-
SEQ ID No. 2209	433.1	Contig53	p	93919	96957	-	-	99%	+	-
SEQ ID No. 2200	432.1	Contig53	m	94806	95345		-	99%		+
SEQ ID No. 2194	431.1	Contig53	m	95892	96227		-	99%		+
SEQ ID No. 104	1071.3	Contig53	m	97055	98368	-	-	98%		-
SEQ ID No. 105	1072.4	Contig53	m	98314	101307	-	-	99%		-
SEQ ID No. 1539	3326.2	Contig53	m	101235	102038	-	-	100%		-
SEQ ID No. 2794	5238.1	Contig53	m	102305	103894	-	-	98%		-
SEQ ID No. 3411	940.4	Contig53	m	103921	104814	-	-	98%		-
SEQ ID No. 3412	941.3	Contig53	m	104792	106147	-	-	100%		-
SEQ ID No. 3413	943.3	Contig53	m	106148	106648	-	-	98%		-
SEQ ID No. 3414	944.3	Contig53	m	106644	107087	-	-	99%		-
SEQ ID No. 1932	3928.1	Contig53	p	107241	109073	-	-	99%		-
SEQ ID No. 1933	3929.2	Contig53	p	109621	111270	+	-	99%		-
SEQ ID No. 1868	3814.1	Contig53	m	111272	112147	+	-	98%		-
SEQ ID No. 3431	969.2	Contig53	p	112570	113343	-	-	99%		-
SEQ ID No. 3430	967.2	Contig53	m	113522	114547	-	-	99%		-

SEQ ID No. 3429	966.2	Contig53	m	114682	115431		-	99%	-
SEQ ID No. 3428	965.3	Contig53	p	115497	116042	+	-	100%	+
SEQ ID No. 1869	3815.2	Contig53	m	116186	116506	-	-	100%	-
SEQ ID No. 1870	3816.2	Contig53	p	116694	117437	+	-	99%	-
SEQ ID No. 1871	3818.1	Contig53	m	117666	117950	-	-	100%	-
SEQ ID No. 1872	3819.2	Contig53	m	118126	118455	-	-	100%	-
SEQ ID No. 2452	4676.2	Contig53	m	118452	119846	-	-	99%	-
SEQ ID No. 3137	6029.1	Contig53	m	120054	120221	-	+		-
SEQ ID No. 2451	4673.1	Contig53	m	120233	120505	-	-	100%	-
SEQ ID No. 2450	4672.1	Contig53	p	120729	121451	-	-	99%	-
SEQ ID No. 2449	4671.3	Contig53	p	121408	122550	-	-	97%	-
SEQ ID No. 2722	5082.3	Contig53	p	122604	124115	-	-	98%	+
SEQ ID No. 141	1123.2	Contig53	p	124321	125466	-	-	99%	-
SEQ ID No. 140	1122.1	Contig53	p	125451	126383	+	-	99%	-
SEQ ID No. 781	2149.4	Contig53	p	126515	127813	-	-	99%	-
SEQ ID No. 2711	5068.3	Contig53	p	128031	128603	-	-	96%	-
SEQ ID No. 2712	5069.3	Contig53	m	128803	129348	-	-	97%	-
SEQ ID No. 3067	5829.2	Contig53	p	129374	130126	+	-	94%	-
SEQ ID No. 2348	4522.4	Contig53	p	130084	130770	-	-	95%	-
SEQ ID No. 2347	4521.1	Contig53	p	130727	131422	-	-	98%	-
SEQ ID No. 2346	4520.1	Contig53	p	131377	132636	-	-	99%	-
SEQ ID No. 362	1479.3	Contig53	p	132452	133864	-	-	97%	-
SEQ ID No. 364	1481.3	Contig53	p	133846	134982	-	-	96%	-
SEQ ID No. 601	1838.2	Contig53	m	135071	136582	+	-	93%	-
SEQ ID No. 2431	4647.2	Contig53	p	136661	137836	+	-	97%	+
SEQ ID No. 2359	4535.2	Contig53	m	137909	139276	+	-	100%	-
SEQ ID No. 2360	4536.1	Contig53	m	139327	140496	+	-	99%	-
SEQ ID No. 2361	4538.2	Contig53	m	140456	142003	-	-	99%	-
SEQ ID No. 2363	4540.2	Contig53	m	142103	142414	+	-	94%	+
SEQ ID No. 735	207.3	Contig53	m	142476	143771	+	-	98%	-
SEQ ID No. 727	206.3	Contig53	p	144091	144813	-	-	95%	-
SEQ ID No. 721	205.1	Contig53	p	144761	145621	+	-	97%	-
SEQ ID No. 717	204.1	Contig53	p	145639	147006	-	-	98%	-
SEQ ID No. 704	202.3	Contig53	p	147129	149489	+	-	99%	-
SEQ ID No. 3264	732.3	Contig53	p	149594	150103	+	-	100%	-
SEQ ID No. 3263	731.2	Contig53	p	150075	151127	-	-	87%	-

SEQ ID No. 3262	730.1	Contig53	p	151252	151707	-	-	97%		-
SEQ ID No. 3261	729.1	Contig53	p	151701	152474	-	-	99%		-
SEQ ID No. 3260	728.2	Contig53	p	152475	152882	-	-	99%		-
SEQ ID No. 1341	3024.1	Contig53	p	152886	154184	-	-	99%		-
SEQ ID No. 1340	3023.1	Contig53	p	154337	154546		+			-
SEQ ID No. 1339	3022.1	Contig53	p	154705	154980	-	-	100%		+
SEQ ID No. 233	1268.2	Contig53	p	155302	156252	-	-	98%		-
SEQ ID No. 232	1267.3	Contig53	p	156234	157100	-	-	97%		-
SEQ ID No. 231	1265.5	Contig53	m	157090	157923	-	-	97%		-
SEQ ID No. 1338	3018.1	Contig53	m	158359	159219	-	-	96%		+
SEQ ID No. 1337	3017.1	Contig53	m	159194	159544		+			+
SEQ ID No. 1336	3016.1	Contig53	m	159545	160066	-	+			+
SEQ ID No. 1335	3015.1	Contig53	m	159996	160766	-	-	97%		-
SEQ ID No. 1334	3014.1	Contig53	m	160720	161361	-	-	97%		-
SEQ ID No. 1333	3013.2	Contig53	m	161325	161795	-	-	96%		-
SEQ ID No. 1332	3011.2	Contig53	m	161789	162874	-	+			-
SEQ ID No. 2814	5288.1	Contig53	m	162784	163212	-	+			-
SEQ ID No. 2815	5289.2	Contig53	m	163191	163865	-	-	94%		-
SEQ ID No. 2824	5312.2	Contig53	p	164165	164791	-	-	96%		+
SEQ ID No. 2825	5313.1	Contig53	m	164821	165249		-	87%		-
SEQ ID No. 2607	4897.2	Contig53	m	165348	166508	-	-	96%		-
SEQ ID No. 2606	4895.1	Contig53	p	166781	167182	+	-	93%		-
SEQ ID No. 2605	4894.1	Contig53	p	167095	167523	-	-	96%		-
SEQ ID No. 671	1945.4	Contig53	p	167507	169294	-	-	99%		-
SEQ ID No. 672	1946.2	Contig53	p	169298	170029	-	-	99%		-
SEQ ID No. 161	1157.4	Contig53	p	170063	172897	-	-	98%		-
SEQ ID No. 1635	3468.2	Contig53	p	172891	174162	-	-	99%		-
SEQ ID No. 416	1559.2	Contig53	p	174213	175388	-	-	99%		-
SEQ ID No. 415	1557.3	Contig53	p	175448	176338	-	-	99%	+	-
SEQ ID No. 905	2356.2	Contig53	p	176752	177408	-	-	100%	+	-
SEQ ID No. 571	1786.2	Contig53	p	177979	178392	-	-	100%		+
SEQ ID No. 570	1784.2	Contig53	m	178463	178897	-	-	100%		+
SEQ ID No. 569	1782.3	Contig53	p	179471	180784	-	-	99%		-
SEQ ID No. 2826	5316.1	Contig53	p	180979	182469	+	-	99%		-
SEQ ID No. 2743	5123.3	Contig53	p	182641	182940	-	-	100%		-
SEQ ID No. 2744	5124.3	Contig53	m	183002	183481	-	+			-

SEQ ID No. 2745	5127.4	Contig53	m	183405	183953	-	-	98%	-
SEQ ID No. 1964	3982.2	Contig53	m	184332	184934	+	-	98%	-
SEQ ID No. 1963	3980.1	Contig53	p	184907	185500	-	-	98%	-
SEQ ID No. 1961	3979.2	Contig53	p	185478	187220	+	-	98%	-
SEQ ID No. 1960	3977.2	Contig53	m	187318	189588	-	-	97%	-
SEQ ID No. 372	1494.2	Contig53	m	189743	190537		-	99%	-
SEQ ID No. 373	1495.1	Contig53	m	190543	190902	-	-	99%	-
SEQ ID No. 374	1497.2	Contig53	m	191202	192272	-	-	98%	-
SEQ ID No. 2827	5317.2	Contig53	m	192709	193332	-	-	97%	-
SEQ ID No. 451	1606.2	Contig53	m	193404	194234	-	-	98%	-
SEQ ID No. 452	1607.2	Contig53	m	194228	194500	-	+		-
SEQ ID No. 453	1608.2	Contig53	p	194570	195826	-	-	99%	-
SEQ ID No. 2396	4597.1	Contig53	m	195867	196286	-	-	100%	+
SEQ ID No. 2397	4598.2	Contig53	p	196513	197271	-	-	99%	-
SEQ ID No. 2490	4729.3	Contig53	p	197390	198196	+	-	100%	-
SEQ ID No. 2489	4727.1	Contig53	p	198224	198724	-	-	99%	-
SEQ ID No. 2488	4726.1	Contig53	p	198889	199224	-	-	98%	+
SEQ ID No. 2487	4725.2	Contig53	m	199470	200573	+	-	99%	-
SEQ ID No. 2486	4724.2	Contig53	p	200568	201215	-	-	100%	-
SEQ ID No. 2485	4721.2	Contig53	p	201197	201745	-	-	99%	-
SEQ ID No. 2593	4878.2	Contig53	m	202045	203496	-	-	99%	-
SEQ ID No. 1228	2859.2	Contig53	p	203633	204940	-	-	99%	-
SEQ ID No. 1227	2858.1	Contig53	m	205210	205464		+		-
SEQ ID No. 3321	81.3	Contig53	m	205742	205948	-	+		-
SEQ ID No. 3314	80.3	Contig53	p	205998	206204	-	+		-
SEQ ID No. 1159	2743.2	Contig53	p	206346	206567	-	+		-
SEQ ID No. 1160	2745.1	Contig53	p	206934	207152	-	+		+
SEQ ID No. 3235	692.1	Contig53	p	209993	210160	-	-	97%	+
SEQ ID No. 3236	693.1	Contig53	m	210303	210845	-	-	96%	-
SEQ ID No. 3237	694.3	Contig53	m	210846	212210	+	-	98%	-
SEQ ID No. 1161	2747.1	Contig53	m	212435	213457	-	-	96%	-
SEQ ID No. 1162	2748.1	Contig53	p	213767	214243	-	-	97%	+
SEQ ID No. 1163	2749.4	Contig53	m	214284	215546	+	-	99%	-
SEQ ID No. 552	1760.4	Contig53	m	215547	216278	-	-	99%	-
SEQ ID No. 550	1759.4	Contig53	p	216391	217173	-	-	99%	+
SEQ ID No. 1720	3591.2	Contig53	p	217448	218038	-	-	100%	-

SEQ ID No. 1721	3592.1	Contig53	p	218007	218450	-	-	99%	+
SEQ ID No. 1722	3593.1	Contig53	p	218407	219393	-	-	98%	-
SEQ ID No. 1723	3594.1	Contig53	m	219507	219902	-	+		+
SEQ ID No. 950	2421.2	Contig53	m	220002	221525	-	-	97%	-
SEQ ID No. 1724	3595.1	Contig53	m	221578	221850	-	-	97%	-
SEQ ID No. 1725	3596.1	Contig53	p	221926	222321	-	-	100%	-
SEQ ID No. 1726	3597.3	Contig53	m	222331	222921	-	-	98%	-
SEQ ID No. 1727	3598.3	Contig53	p	223164	223358	-	+		+
SEQ ID No. 1728	3599.4	Contig53	p	223351	224187	-	-	99%	-
SEQ ID No. 3039	5741.1	Contig53	m	224214	224969	-	-	99%	-
SEQ ID No. 1106	2657.3	Contig53	p	225046	227253	-	-	98%	-
SEQ ID No. 1105	2655.1	Contig53	p	227254	227673	-	-	97%	-
SEQ ID No. 1104	2654.1	Contig53	p	227495	227683		-	90%	+
SEQ ID No. 1103	2653.1	Contig53	m	227848	228192		-	93%	+
SEQ ID No. 2559	482.2	Contig53	m	228193	230004	-	-	98%	-
SEQ ID No. 2550	481.1	Contig53	p	230355	230837	-	-	99%	-
SEQ ID No. 2541	480.3	Contig53	p	230827	232278	-	-	99%	-
SEQ ID No. 677	1951.4	Contig53	p	232262	233032	-	-	100%	-
SEQ ID No. 676	1950.2	Contig53	p	233014	234315	-	-	100%	-
SEQ ID No. 1606	342.2	Contig53	p	234284	235549	-	-	100%	-
SEQ ID No. 1612	343.1	Contig53	p	235534	235995	-	-	100%	-
SEQ ID No. 1616	344.1	Contig53	p	236025	236402	-	-	100%	-
SEQ ID No. 1630	346.2	Contig53	p	236327	237358	-	-	100%	-
SEQ ID No. 1636	347.1	Contig53	m	237431	238357	-	-	100%	-
SEQ ID No. 1649	349.2	Contig53	m	238592	239380	-	-	100%	-
SEQ ID No. 1102	2652.2	Contig53	m	239239	240552	-	-	100%	-
SEQ ID No. 1101	2650.1	Contig53	p	241173	242501	+	-	100%	-
SEQ ID No. 1021	2521.4	Contig53	m	242561	243616	-	-	100%	-
SEQ ID No. 1022	2522.1	Contig53	m	243808	244155	+	-	100%	+
SEQ ID No. 1023	2523.1	Contig53	p	244324	244743	-	-	100%	-
SEQ ID No. 1024	2524.2	Contig53	m	244779	245492	-	-	97%	-
SEQ ID No. 1100	2649.1	Contig53	p	245702	246898	-	+		-
SEQ ID No. 3210	650.3	Contig53	p	246861	248021	-	+		-
SEQ ID No. 3211	651.2	Contig53	m	248091	248534	-	+		-
SEQ ID No. 3212	652.1	Contig53	p	248604	249497	-	-	96%	-
SEQ ID No. 1099	2648.2	Contig53	p	249860	250981	+	-	99%	-

SEQ ID No. 1098	2647.2	Contig53	m	251141	251743	-	-	100%	-
SEQ ID No. 1097	2646.1	Contig53	p	251907	252410	-	-	100%	-
SEQ ID No. 1096	2645.1	Contig53	m	252511	253491	-	-	100%	-
SEQ ID No. 248	1288.2	Contig53	m	253575	255041	-	-	100%	-
SEQ ID No. 655	1915.3	Contig53	m	255160	257109	-	-	100%	-
SEQ ID No. 654	1913.2	Contig53	m	257214	257630	-	-	100%	-
SEQ ID No. 1095	2644.1	Contig53	p	257737	258120	+	-	100%	+
SEQ ID No. 1094	2643.1	Contig53	p	258216	259628	-	-	100%	-
SEQ ID No. 339	1434.3	Contig53	m	259654	261867	-	-	100%	-
SEQ ID No. 697	2000.1	Contig53	m	261939	262439	-	-	100%	+
SEQ ID No. 1249	290.2	Contig53	m	262400	265921	+	-	100%	-
SEQ ID No. 1229	286.2	Contig53	m	265922	266464	+	-	100%	-
SEQ ID No. 1221	284.2	Contig53	m	266431	267564	-	-	100%	-
SEQ ID No. 1215	283.1	Contig53	m	267495	268037	-	-	100%	-
SEQ ID No. 1208	282.2	Contig53	m	268047	268784	-	-	100%	-
SEQ ID No. 632	1882.2	Contig53	m	268976	269923		-	100%	-
SEQ ID No. 633	1883.4	Contig53	m	270013	271392	-	-	100%	+
SEQ ID No. 2759	5159.2	Contig53	p	271638	273131		-	100%	-
SEQ ID No. 660	1926.2	Contig53	p	273283	273951	-	-	98%	-
SEQ ID No. 661	1928.2	Contig53	p	273981	275270	-	-	99%	-
SEQ ID No. 3322	813.2	Contig53	p	275271	276566	-	-	100%	-
SEQ ID No. 3323	815.2	Contig53	p	276554	277786	-	-	97%	-
SEQ ID No. 1895	3860.2	Contig53	p	277948	278538	-	-	100%	-
SEQ ID No. 1894	3859.2	Contig53	p	278522	279865	-	-	99%	-
SEQ ID No. 1127	2690.2	Contig53	p	280065	281204	-	-	97%	+
SEQ ID No. 1125	2688.2	Contig53	p	281205	281639	-	+		+
SEQ ID No. 2758	5156.1	Contig53	m	281789	282058	-	+		+
SEQ ID No. 694	1998.3	Contig53	m	282043	283029	-	-	98%	-
SEQ ID No. 3037	5739.3	Contig53	p	283472	288685	-	+		-
SEQ ID No. 3410	94.7	Contig53	p	288516	295643	-	+		+
SEQ ID No. 1085	2624.1	Contig54	p	3	554		+		-
SEQ ID No. 57	10.1	Contig54	p	888	1235		+		-
SEQ ID No. 3382	9.1	Contig54	p	1049	1552		+		-
SEQ ID No. 3242	7.1	Contig54	p	1711	5829	-	+		-
SEQ ID No. 1977	4.1	Contig54	p	5766	6818	-	+		-
SEQ ID No. 1322	3.1	Contig54	p	6822	7124		+		-

SEQ ID No. 695	2.1	Contig54	p	7066	7698	-	+		-
SEQ ID No. 3194	63.1	Contig54	m	9066	9281	-	+		+
SEQ ID No. 3167	61.1	Contig54	p	9759	10058		+		+
SEQ ID No. 3126	60.1	Contig54	m	9838	10008	-	+		+
SEQ ID No. 3091	59.1	Contig54	p	10665	11108		+		-
SEQ ID No. 3052	58.1	Contig54	m	11098	11802	-	+		-
SEQ ID No. 3025	57.1	Contig54	p	11919	12803	-	-	91%	+
SEQ ID No. 2919	55.1	Contig54	p	12574	13131	-	-	96%	+
SEQ ID No. 2775	52.1	Contig54	p	13132	13347	-	+		-
SEQ ID No. 2666	50.1	Contig54	p	13299	13646	+	-	82%	-
SEQ ID No. 2540	48.1	Contig54	p	13647	13937	-	+		-
SEQ ID No. 2466	47.1	Contig54	p	13831	16404	-	-	94%	-
SEQ ID No. 2399	46.1	Contig54	p	16389	17111	+	-	93%	-
SEQ ID No. 2333	45.1	Contig54	p	17248	17664	-	-	80%	+
SEQ ID No. 2266	44.1	Contig54	p	17665	18708	-	-	94%	-
SEQ ID No. 2187	43.1	Contig54	p	18684	19421	-	-	95%	-
SEQ ID No. 2122	42.1	Contig54	p	19373	20170	+	-	95%	-
SEQ ID No. 1914	39.1	Contig54	p	20125	21258	-	-	93%	-
SEQ ID No. 1801	37.1	Contig54	p	21218	22264	-	-	91%	-
SEQ ID No. 1729	36.1	Contig54	p	22191	24158	-	-	98%	-
SEQ ID No. 1654	35.1	Contig54	p	24187	25053	+	-	98%	+
SEQ ID No. 1592	34.1	Contig54	p	25266	26048	-	+		+
SEQ ID No. 1522	33.1	Contig54	m	26172	28868	-	+		-
SEQ ID No. 1323	30.1	Contig54	p	28991	29266	-	+		+
SEQ ID No. 1248	29.1	Contig54	p	29232	29522		+		-
SEQ ID No. 1131	27.1	Contig54	p	29610	30263		+		-
SEQ ID No. 1072	26.1	Contig54	p	30264	31076	-	+		-
SEQ ID No. 1007	25.1	Contig54	m	31103	31582	-	+		-
SEQ ID No. 933	24.1	Contig54	m	31557	33023		+		-
SEQ ID No. 696	20.1	Contig54	m	32977	33345	-	+		+
SEQ ID No. 644	19.1	Contig54	m	33353	34066	-	+		-
SEQ ID No. 580	18.1	Contig54	p	34124	34408		+		+
SEQ ID No. 446	16.1	Contig54	p	34418	34645	-	+		-
SEQ ID No. 375	15.1	Contig54	p	34591	35802	-	+		-
SEQ ID No. 254	13.1	Contig54	p	35966	36853	-	+		+
SEQ ID No. 126	11.2	Contig54	p	36783	37802	-	+		-

SEQ ID No. 1084	2622.1	Contig54	p	38133	40073	-	+		-
SEQ ID No. 1076	2605.3	Contig54	p	40025	42568	-	+		-
SEQ ID No. 1083	2620.1	Contig54	p	42674	43009	-	-	97%	-
SEQ ID No. 1081	2619.1	Contig54	p	42970	43425		+		-
SEQ ID No. 1080	2616.1	Contig54	p	43401	44285		+		-
SEQ ID No. 308	139.5	Contig54	p	44404	48882	-	+		-
SEQ ID No. 1079	2611.1	Contig54	p	49022	49240		-	83%	+
SEQ ID No. 1078	2609.2	Contig54	p	49931	50938	-	-	94%	+
SEQ ID No. 1077	2607.2	Contig54	p	51047	52288	+	-	98%	-
SEQ ID No. 1648	3489.2	Contig54	m	52297	53340	-	-	94%	+
SEQ ID No. 432	1582.2	Contig54	p	53466	53990	-	-	98%	-
SEQ ID No. 431	1581.4	Contig54	m	54089	55366	-	-	87%	+
SEQ ID No. 160	1156.3	Contig54	m	55687	56997		+		-
SEQ ID No. 159	1153.1	Contig54	m	57180	57608		-	98%	-
SEQ ID No. 158	1152.2	Contig54	m	57786	58682	-	-	92%	-
SEQ ID No. 1647	3488.1	Contig54	m	58908	61016	-	-	86%	+
SEQ ID No. 1646	3486.1	Contig54	m	61261	61521		+		+
SEQ ID No. 1645	3485.1	Contig54	p	61495	62223	-	-	87%	+
SEQ ID No. 1644	3484.1	Contig54	m	62317	62967	-	-	93%	+
SEQ ID No. 1643	3483.3	Contig54	m	63171	63917	-	-	92%	-
SEQ ID No. 588	1809.4	Contig54	m	64253	66679	-	-	84%	-
SEQ ID No. 1642	3481.2	Contig54	m	66646	66918	-	-	100%	-
SEQ ID No. 1641	3480.4	Contig54	m	66923	68722	-	-	98%	-
SEQ ID No. 3334	833.3	Contig54	p	68959	71343	+	-	93%	-
SEQ ID No. 757	2116.2	Contig54	m	71452	71955	-	-	96%	-
SEQ ID No. 2659	499.2	Contig54	m	72017	73159		-	97%	-
SEQ ID No. 2667	500.2	Contig54	p	73382	74464	+	-	98%	-
SEQ ID No. 2672	501.2	Contig54	p	74623	75396	-	-	96%	+
SEQ ID No. 2677	502.3	Contig54	m	75598	76065	-	+		+
SEQ ID No. 2682	503.3	Contig54	p	76245	77588		-	98%	-
SEQ ID No. 1435	3161.1	Contig54	m	77714	78490	-	-	99%	-
SEQ ID No. 1434	3160.1	Contig54	m	78711	79331	-	-	87%	+
SEQ ID No. 90	1049.2	Contig54	p	79487	80278	-	-	96%	-
SEQ ID No. 91	1050.4	Contig54	p	80420	83893	-	-	98%	+
SEQ ID No. 915	2368.5	Contig54	m	83988	85529	-	-	94%	-
SEQ ID No. 564	1778.5	Contig54	p	85555	86043	-	-	100%	-

SEQ ID No. 563	1775.3	Contig54	p	86468	88030	-	-	94%	+
SEQ ID No. 1860	3801.2	Contig54	m	88195	89649	+	-	98%	-
SEQ ID No. 1859	3800.4	Contig54	m	89808	90146	-	-	100%	+
SEQ ID No. 3307	791.5	Contig54	m	90299	91600	-	-	97%	-
SEQ ID No. 1858	3798.1	Contig54	m	91480	91875		-	99%	+
SEQ ID No. 1857	3797.1	Contig54	m	92077	93171	-	-	96%	-
SEQ ID No. 1856	3796.3	Contig54	m	93107	93736	-	-	99%	-
SEQ ID No. 1855	3793.3	Contig54	p	93867	94790	+	-	94%	-
SEQ ID No. 1854	3792.1	Contig54	m	94897	95136	+	+		+
SEQ ID No. 1853	3791.1	Contig54	m	95344	96318	-	-	94%	+
SEQ ID No. 1852	3789.2	Contig54	m	96578	97606	-	-	98%	-
SEQ ID No. 1851	3788.1	Contig54	m	97597	98373	-	-	99%	-
SEQ ID No. 1850	3785.2	Contig54	m	98361	99386	-	-	99%	-
SEQ ID No. 1849	3784.2	Contig54	m	99175	100311	-	-	98%	-
SEQ ID No. 1848	3783.1	Contig54	m	100438	100734	-	-	94%	-
SEQ ID No. 904	2355.2	Contig54	m	100931	101824	-	-	97%	-
SEQ ID No. 1847	3780.2	Contig54	m	102158	102685	-	-	95%	-
SEQ ID No. 379	1505.4	Contig54	p	102899	105229		-	96%	+
SEQ ID No. 380	1507.3	Contig54	p	105338	106840	-	-	98%	-
SEQ ID No. 1626	3454.1	Contig54	p	106841	107404	-	-	91%	+
SEQ ID No. 1625	3453.1	Contig54	p	107695	109209	-	-	90%	-
SEQ ID No. 1624	3451.1	Contig54	p	109351	110169	-	-	95%	+
SEQ ID No. 1623	3450.2	Contig54	p	110251	111336	-	-	98%	-
SEQ ID No. 2750	5135.2	Contig54	m	111413	112909	-	-	93%	+
SEQ ID No. 1035	2545.2	Contig54	m	113096	114103	-	-	90%	+
SEQ ID No. 1036	2546.4	Contig54	m	113946	115451		-	94%	-
SEQ ID No. 1935	3933.3	Contig54	p	115898	117001	-	-	96%	+
SEQ ID No. 229	1261.2	Contig54	p	117232	118920	-	-	99%	-
SEQ ID No. 230	1264.3	Contig54	m	118978	120174	+	-	95%	-
SEQ ID No. 2322	4485.2	Contig54	p	120515	121345	+	-	98%	-
SEQ ID No. 2321	4484.1	Contig54	p	121446	122228	-	-	97%	+
SEQ ID No. 3378	895.3	Contig54	p	122576	123634	-	-	98%	-
SEQ ID No. 3377	894.3	Contig54	m	123736	124005	-	-	98%	+
SEQ ID No. 823	2220.2	Contig54	p	124301	124963	-	-	99%	-
SEQ ID No. 821	2219.2	Contig54	m	124648	125064		-	98%	-
SEQ ID No. 2199	4319.2	Contig54	m	125030	125764	+	-	98%	+

SEQ ID No. 144	1126.3	Contig54	p	125926	127218	-	-	96%		-
SEQ ID No. 143	1125.2	Contig54	p	127188	128384	-	-	97%		-
SEQ ID No. 142	1124.2	Contig54	p	128626	129426	-	-	97%		-
SEQ ID No. 2198	4316.4	Contig54	m	129491	131527	-	-	97%		+
SEQ ID No. 2857	5381.1	Contig54	p	129885	130109		-	100%		+
SEQ ID No. 2338	4508.4	Contig54	p	131646	132248	-	-	98%		-
SEQ ID No. 129	1105.2	Contig54	m	132369	134687	-	-	96%		-
SEQ ID No. 2129	4211.2	Contig54	m	134827	136197	-	-	92%		-
SEQ ID No. 2130	4212.1	Contig54	p	136221	137318	-	-	97%		-
SEQ ID No. 151	1137.2	Contig54	p	137371	137796	-	-	98%		+
SEQ ID No. 150	1136.3	Contig54	p	137907	139430	+	-	99%		-
SEQ ID No. 2131	4214.3	Contig54	p	139523	141115	-	-	96%		-
SEQ ID No. 2886	544.2	Contig54	p	141103	142440	-	-	96%		-
SEQ ID No. 2883	543.1	Contig54	p	142846	143115	-	-	100%		+
SEQ ID No. 985	2468.2	Contig54	m	143553	144497	-	-	97%		-
SEQ ID No. 986	2469.2	Contig54	m	144767	146230		-	99%		-
SEQ ID No. 988	2471.3	Contig54	p	146263	147339	-	-	98%		-
SEQ ID No. 599	1832.4	Contig54	p	147324	147938	-	-	98%		-
SEQ ID No. 598	1831.3	Contig54	p	147869	149143	-	-	97%		-
SEQ ID No. 1815	3721.1	Contig54	p	149133	149618	-	-	100%		-
SEQ ID No. 1633	3465.2	Contig54	p	149702	151381	+	-	100%		-
SEQ ID No. 1632	3463.1	Contig54	p	151330	152202	-	-	100%		-
SEQ ID No. 1631	3461.4	Contig54	p	152359	154266	-	-	97%		+
SEQ ID No. 865	2293.5	Contig54	p	154352	156745	-	-	97%		-
SEQ ID No. 322	1409.3	Contig54	p	157120	157902	+	-	99%	+	-
SEQ ID No. 324	1410.2	Contig54	m	158050	158997	-	-	98%	+	-
SEQ ID No. 1629	3459.3	Contig54	m	159106	160590		-	98%		-
SEQ ID No. 1628	3456.4	Contig54	m	160503	161009	+	+			-
SEQ ID No. 1627	3455.3	Contig54	m	161013	162041	-	-	97%		+
SEQ ID No. 449	1602.2	Contig54	p	162071	163246	-	-	97%		-
SEQ ID No. 448	1601.3	Contig54	p	163332	164519	+	-	98%		-
SEQ ID No. 2813	5282.3	Contig54	p	164626	164874	-	+			-
SEQ ID No. 3135	6019.1	Contig54	p	164875	165402		-	98%		-
SEQ ID No. 1844	3771.2	Contig54	m	165491	166123	-	-	99%		-
SEQ ID No. 3347	854.3	Contig54	m	166111	166785	-	-	96%		-
SEQ ID No. 3050	5780.1	Contig54	m	166583	166912	-	+			-

SEQ ID No. 3346	853.1	Contig54	m	166885	167625	-	-	98%		-
SEQ ID No. 3345	852.3	Contig54	m	167598	168224	-	-	98%		-
SEQ ID No. 1845	3772.2	Contig54	m	168194	169255	-	-	97%		+
SEQ ID No. 127	1101.4	Contig54	m	169230	170324	-	-	97%		-
SEQ ID No. 128	1102.2	Contig54	m	170325	171641	-	-	96%		-
SEQ ID No. 843	2255.2	Contig54	m	171626	172522	-	-	96%		-
SEQ ID No. 1846	3778.2	Contig54	m	172504	172809	-	-	98%		-
SEQ ID No. 2947	5550.3	Contig54	p	173129	174709	-	-	98%		-
SEQ ID No. 2948	5552.2	Contig54	p	174713	175861	-	-	99%		-
SEQ ID No. 2275	4409.1	Contig54	m	176087	176728	-	-	100%		-
SEQ ID No. 2274	4408.1	Contig54	m	176920	177180		-	98%		+
SEQ ID No. 2273	4407.1	Contig54	p	176950	177201	-	+			-
SEQ ID No. 2272	4406.1	Contig54	p	177700	178284	-	-	99%		-
SEQ ID No. 2271	4405.1	Contig54	p	178381	178899	-	-	98%		-
SEQ ID No. 2270	4404.1	Contig54	m	178929	179393	-	-	94%		-
SEQ ID No. 2269	4403.2	Contig54	p	179417	180214	-	-	91%		-
SEQ ID No. 2268	4402.2	Contig54	p	180142	180759	-	-	92%		-
SEQ ID No. 1070	2598.3	Contig54	m	180883	181581	+	-	97%		-
SEQ ID No. 2741	5115.2	Contig54	p	181805	182791	-	-	99%		-
SEQ ID No. 2742	5116.3	Contig54	p	182843	183412	-	-	97%		-
SEQ ID No. 2949	5554.2	Contig54	p	183792	184052		+			+
SEQ ID No. 3134	6018.1	Contig54	m	184053	184367	-	+			+
SEQ ID No. 2785	5217.1	Contig55	p	45	509	-	+			-
SEQ ID No. 1107	2659.2	Contig55	p	502	996		+			-
SEQ ID No. 1108	2660.1	Contig55	p	983	1276	-	+			-
SEQ ID No. 1109	2661.2	Contig55	p	1221	1838	-	+			+
SEQ ID No. 1110	2662.2	Contig55	p	1823	2554	+	+		+	-
SEQ ID No. 2574	485.3	Contig55	p	2536	4023	+	+		+	+
SEQ ID No. 2580	486.2	Contig55	p	3956	4243	-	+			+
SEQ ID No. 1032	254.2	Contig55	p	4093	6849	-	+		+	-
SEQ ID No. 1038	255.1	Contig55	p	6821	7174	+	+			-
SEQ ID No. 1045	256.1	Contig55	p	7135	7797	-	+			-
SEQ ID No. 1058	258.2	Contig55	p	7683	8804		+			-
SEQ ID No. 698	2006.2	Contig55	p	8818	9483	+	+			-
SEQ ID No. 280	1344.3	Contig55	p	9401	10501	+	+			-
SEQ ID No. 2786	5219.1	Contig55	p	10464	11285	-	+			-

SEQ ID No. 279	1342.3	Contig55	p	11279	12070	+	+			-
SEQ ID No. 722	2051.2	Contig55	p	12067	12564		+			-
SEQ ID No. 253	1299.3	Contig55	p	12561	13958	+	+		+	-
SEQ ID No. 3010	567.5	Contig55	p	13877	16723	-	+		+	+
SEQ ID No. 818	2212.4	Contig55	p	16754	18805	+	+		+	-
SEQ ID No. 445	1599.4	Contig55	p	18806	24001	-	+		+	-
SEQ ID No. 2787	5220.1	Contig55	p	23872	24702	-	+			-
SEQ ID No. 1111	2665.1	Contig55	p	24873	25163	-	+			+
SEQ ID No. 1112	2666.1	Contig55	m	25296	25622	-	+			+
SEQ ID No. 1113	2667.1	Contig55	m	25607	26068	-	+			+
SEQ ID No. 3255	717.2	Contig55	p	26412	26885	-	+			+
SEQ ID No. 3254	716.1	Contig55	p	26900	27130	-	+			+
SEQ ID No. 3253	715.1	Contig55	p	27284	27616		+			+
SEQ ID No. 3252	713.2	Contig55	p	27686	28033		+			+
SEQ ID No. 3069	5844.1	Contig55	m	28321	28734	-	+			-
SEQ ID No. 2320	4483.2	Contig55	m	28848	29474		+			+
SEQ ID No. 2319	4482.1	Contig55	p	29686	30615	-	+			-
SEQ ID No. 2318	4481.1	Contig55	m	30697	31317	-	+			+
SEQ ID No. 1752	3632.3	Contig55	m	31318	34410	-	+			-
SEQ ID No. 836	2244.2	Contig55	p	34509	35120	-	+			-
SEQ ID No. 837	2245.2	Contig55	p	35114	36196	-	+			-
SEQ ID No. 838	2247.5	Contig55	p	36168	37979	-	+			-
SEQ ID No. 3301	783.2	Contig55	m	38018	38377	-	+			-
SEQ ID No. 3302	785.4	Contig55	m	38404	39363		+			+
SEQ ID No. 3070	5846.2	Contig55	m	39236	39436	-	+			+
SEQ ID No. 1751	3631.4	Contig55	m	39466	39723	-	+			+
SEQ ID No. 3133	6016.1	Contig55	m	39777	40088	-	+			-
SEQ ID No. 3132	6015.1	Contig55	p	40166	40540		+			+
SEQ ID No. 2739	5113.3	Contig55	p	40350	40982		+			+
SEQ ID No. 2740	5114.3	Contig55	p	40999	41910		+			+
SEQ ID No. 1420	3138.2	Contig55	p	41939	42838		+			+
SEQ ID No. 1419	3137.1	Contig55	p	42851	43375	+	+			+
SEQ ID No. 1418	3136.1	Contig55	p	43426	43785	-	+			-
SEQ ID No. 3313	799.3	Contig55	p	43996	45474	-	+		+	-
SEQ ID No. 3312	798.2	Contig55	p	45371	47254	+	+		+	-
SEQ ID No. 3311	796.2	Contig55	p	47244	48029	-	+		+	-

SEQ ID No. 745	2087.2	Contig55	p	48030	48989	-	+	+	-
SEQ ID No. 803	2180.2	Contig55	p	48956	50470	-	+		-
SEQ ID No. 1417	3135.1	Contig55	p	50443	50679	-	+		-
SEQ ID No. 1416	3134.1	Contig55	p	50672	51244	-	+		-
SEQ ID No. 554	1762.2	Contig55	p	51309	52265	-	+		-
SEQ ID No. 553	1761.4	Contig55	p	52416	55592	-	+		-
SEQ ID No. 426	1575.2	Contig55	p	55896	56345	-	+		-
SEQ ID No. 1553	3346.2	Contig55	p	56297	58270	-	+		-
SEQ ID No. 1554	3347.1	Contig55	p	58332	59138	+	+		-
SEQ ID No. 1555	3348.2	Contig55	p	59139	60737		+		-
SEQ ID No. 1556	3349.3	Contig55	p	60738	61637	-	+		-
SEQ ID No. 1558	3350.3	Contig55	p	61771	62142	-	+		-
SEQ ID No. 1559	3351.3	Contig55	p	62162	63334	-	+	+	-
SEQ ID No. 1560	3352.1	Contig55	m	63363	63833	-	+		-
SEQ ID No. 1561	3353.1	Contig55	p	63922	65490	-	+	+	-
SEQ ID No. 1562	3354.1	Contig55	p	65511	66116	-	+	+	-
SEQ ID No. 3426	960.2	Contig55	p	66059	66436	-	+		-
SEQ ID No. 3424	959.2	Contig55	p	66489	66674		+		+
SEQ ID No. 3423	957.2	Contig55	m	66768	67700	-	+		-
SEQ ID No. 3422	956.2	Contig55	m	67701	68516	-	+		-
SEQ ID No. 1563	3355.2	Contig55	m	68675	69007	-	+		-
SEQ ID No. 1564	3356.2	Contig55	m	68952	69212	-	+		-
SEQ ID No. 1565	3357.2	Contig55	m	69452	69934	-	+		+
SEQ ID No. 2757	5154.2	Contig55	m	70538	71368	-	+		-
SEQ ID No. 2756	5153.1	Contig55	m	71343	71813	-	+		+
SEQ ID No. 1054	2573.3	Contig55	m	71767	72372	-	+		-
SEQ ID No. 1241	2881.2	Contig55	m	72633	73565	-	+		+
SEQ ID No. 1240	2879.3	Contig55	m	73627	74961	-	+		+
SEQ ID No. 1239	2878.2	Contig55	p	75141	75842	-	+		-
SEQ ID No. 290	1362.3	Contig55	p	76079	77905	-	+		+
SEQ ID No. 291	1364.2	Contig55	m	77909	79549	+	+		-
SEQ ID No. 1238	2877.2	Contig55	m	79612	80280	-	+		-
SEQ ID No. 800	2175.3	Contig55	m	80356	80577	-	+		+
SEQ ID No. 799	2174.1	Contig55	m	80578	80886	-	+		-
SEQ ID No. 110	1078.2	Contig55	p	81022	82008	-	+		+
SEQ ID No. 109	1077.1	Contig55	p	82086	82808	-	+		-

SEQ ID No. 108	1076.3	Contig55	p	82869	83252	-	+			-
SEQ ID No. 798	2173.2	Contig55	p	83159	83683	-	+			-
SEQ ID No. 797	2172.2	Contig55	p	83680	84168	-	+			-
SEQ ID No. 1237	2874.1	Contig55	p	84158	84934		+			-
SEQ ID No. 1236	2873.1	Contig55	m	84965	85960	-	+			+
SEQ ID No. 3285	761.2	Contig55	m	85953	86924	-	+		+	-
SEQ ID No. 3286	762.2	Contig55	p	86954	87676	+	+		+	-
SEQ ID No. 3287	764.1	Contig55	p	87747	88268	-	+		+	-
SEQ ID No. 3288	765.1	Contig55	m	88340	89080	-	+			-
SEQ ID No. 3289	766.3	Contig55	m	89231	90691	-	+		+	+
SEQ ID No. 3371	886.3	Contig55	m	90736	91527	-	+		+	-
SEQ ID No. 3370	885.2	Contig55	p	91494	93173	-	+			-
SEQ ID No. 3369	884.3	Contig55	p	93167	94135	-	+			+
SEQ ID No. 2367	4549.2	Contig55	p	94210	95664	-	+			-
SEQ ID No. 213	1237.2	Contig55	p	95677	96117	-	+			-
SEQ ID No. 212	1236.2	Contig55	m	96236	97093	-	-	96%		-
SEQ ID No. 211	1235.3	Contig55	m	97094	97873	-	+			-
SEQ ID No. 2755	5152.1	Contig55	m	97828	98214	-	+			-
SEQ ID No. 2754	5151.1	Contig55	m	98262	99167	-	+			-
SEQ ID No. 610	1848.4	Contig55	m	99168	100433	-	+			-
SEQ ID No. 611	1849.3	Contig55	m	100339	100893	-	+			-
SEQ ID No. 613	1852.2	Contig55	m	100904	101302	-	+		+	-
SEQ ID No. 392	1522.2	Contig55	m	101257	102432	-	+			-
SEQ ID No. 393	1523.2	Contig55	m	102483	102815	-	+			+
SEQ ID No. 394	1524.3	Contig55	m	102816	103892	-	+			-
SEQ ID No. 844	2258.2	Contig55	m	104112	105362	-	+			+
SEQ ID No. 846	2260.2	Contig55	m	105337	106014	-	+		+	-
SEQ ID No. 1344	3030.2	Contig55	m	106015	106557	-	+			-
SEQ ID No. 1345	3031.2	Contig55	m	106523	106960	-	+			-
SEQ ID No. 1346	3035.2	Contig55	p	107019	107624	-	+			-
SEQ ID No. 3215	659.2	Contig55	m	107629	108609	-	+			-
SEQ ID No. 3216	660.1	Contig55	m	108579	109169	-	+			-
SEQ ID No. 3217	661.3	Contig55	m	109132	110268	-	+			-
SEQ ID No. 1347	3036.2	Contig55	m	110272	110826	-	+			-
SEQ ID No. 801	2176.3	Contig55	m	110897	111364	-	+			-
SEQ ID No. 802	2177.2	Contig55	p	111374	112039	-	+			-

SEQ ID No. 3265	737.1	Contig55	m	112373	113413	-	+			-
SEQ ID No. 3266	738.2	Contig55	m	113352	113906	-	+			-
SEQ ID No. 3267	739.4	Contig55	m	113977	114864	-	+			-
SEQ ID No. 2634	4947.2	Contig55	p	115163	116416	-	+			-
SEQ ID No. 2633	4945.2	Contig55	p	116406	117425	-	+			-
SEQ ID No. 2719	5078.2	Contig55	p	117433	118752	-	+			-
SEQ ID No. 2721	5081.3	Contig55	p	118891	119469	+	+		+	-
SEQ ID No. 2720	5080.4	Contig55	p	119493	120956	-	+			-
SEQ ID No. 2737	5106.3	Contig55	m	120958	122001	-	+			-
SEQ ID No. 2738	5108.2	Contig55	p	122110	122373		+			+
SEQ ID No. 2753	5147.1	Contig55	p	122374	123642	-	+			+
SEQ ID No. 1656	3500.3	Contig55	p	123593	124777	-	+			+
SEQ ID No. 1653	3499.1	Contig55	p	124882	125178		+			+
SEQ ID No. 205	1226.2	Contig55	p	125159	126205	-	+			+
SEQ ID No. 204	1225.2	Contig55	p	126202	127263	-	+			-
SEQ ID No. 856	2275.3	Contig55	p	127253	128290	-	+		+	-
SEQ ID No. 857	2277.2	Contig55	p	128618	128899	-	+		+	-
SEQ ID No. 171	1171.3	Contig55	p	128900	129670	+	+			-
SEQ ID No. 172	1172.2	Contig55	p	129671	130108	-	+			-
SEQ ID No. 1652	3498.1	Contig55	p	130156	130584	+	+			+
SEQ ID No. 3186	616.4	Contig55	p	130624	131685	-	+			+
SEQ ID No. 3100	591.3	Contig55	p	131637	132476	-	+			-
SEQ ID No. 2834	5328.1	Contig55	m	132518	133048	-	+			-
SEQ ID No. 2127	4209.2	Contig55	m	133049	134221	-	+			-
SEQ ID No. 868	2298.2	Contig55	m	134262	135566	-	-			+
SEQ ID No. 79	1035.3	Contig55	p	136010	137455	-	-	95%		-
SEQ ID No. 2311	4473.4	Contig55	m	137527	138237	-	-	97%		-
SEQ ID No. 2312	4474.1	Contig55	m	138206	138715	-	-	99%		-
SEQ ID No. 2313	4475.1	Contig55	p	138772	139452		-	100%		-
SEQ ID No. 2314	4477.2	Contig55	p	139582	142005	-	-	98%		-
SEQ ID No. 2315	4479.2	Contig55	p	142158	142640	-	-	98%		+
SEQ ID No. 2317	4480.4	Contig55	m	142691	143992	-	-	99%		-
SEQ ID No. 2727	5090.2	Contig55	m	143953	145032	-	-	97%		-
SEQ ID No. 2725	5088.2	Contig55	m	145313	145738	-	+	99%		-
SEQ ID No. 3006	5660.1	Contig55	p	145939	146925	-	+			-
SEQ ID No. 2701	5056.2	Contig55	p	147059	147427	-	+			-

SEQ ID No. 2702	5058.2	Contig55	m	147683	147949	-	+		-
SEQ ID No. 2703	5059.3	Contig55	m	147915	148292	-	+		-
SEQ ID No. 3136	602.4	Contig55	m	148411	149406	-	-	98%	-
SEQ ID No. 3142	604.3	Contig55	m	149566	150327	-	-	96%	-
SEQ ID No. 2414	4623.2	Contig55	m	150381	150938	-	-	94%	-
SEQ ID No. 2700	5054.2	Contig55	m	150905	151864	-	-	95%	-
SEQ ID No. 1982	4004.2	Contig55	m	151865	153652	-	-	98%	-
SEQ ID No. 1981	4003.1	Contig55	m	153639	155102	-	-	96%	-
SEQ ID No. 202	1223.2	Contig55	p	155439	156341		-	80%	-
SEQ ID No. 203	1224.2	Contig55	m	156401	157246	+	-	97%	-
SEQ ID No. 1980	4002.2	Contig55	m	157251	159182	+	-	95%	-
SEQ ID No. 1979	4000.2	Contig55	m	159161	159628	-	-	97%	-
SEQ ID No. 2192	4307.2	Contig55	m	159772	161094	-	-	95%	+
SEQ ID No. 847	2261.3	Contig55	m	161261	162034	-	-	93%	+
SEQ ID No. 124	1098.3	Contig55	m	162153	163196	-	-	97%	-
SEQ ID No. 125	1099.2	Contig55	p	163387	164445	-	-	94%	-
SEQ ID No. 2191	4305.2	Contig55	m	164470	164892	-	-	90%	+
SEQ ID No. 2962	5582.4	Contig55	m	165100	167310	-	-	94%	-
SEQ ID No. 2041	4088.2	Contig55	p	167432	168376	-	-	97%	-
SEQ ID No. 2040	4087.2	Contig55	m	168470	169795	-	-	97%	-
SEQ ID No. 134	1111.3	Contig55	m	169738	170574	-	-	98%	-
SEQ ID No. 132	1109.3	Contig55	m	170559	171554	-	-	97%	-
SEQ ID No. 131	1107.2	Contig55	m	171541	172536	-	-	97%	-
SEQ ID No. 2039	4084.2	Contig55	p	172650	173840	-	-	97%	-
SEQ ID No. 2038	4083.3	Contig55	p	173810	174493	-	-	98%	+
SEQ ID No. 427	1577.5	Contig55	m	174673	177582	-	-	95%	-
SEQ ID No. 2210	4332.1	Contig55	p	177722	178267	-	-	98%	-
SEQ ID No. 2211	4333.1	Contig55	p	178294	179169	-	-	95%	-
SEQ ID No. 2212	4334.1	Contig55	p	179170	180063	-	-	98%	-
SEQ ID No. 2213	4336.2	Contig55	p	180168	181493	-	-	100%	-
SEQ ID No. 2421	4632.1	Contig55	m	181506	182885	-	-	99%	+
SEQ ID No. 2422	4633.2	Contig55	p	182895	183470	-	-	100%	-
SEQ ID No. 2423	4634.2	Contig55	m	183488	185485	-	-	97%	-
SEQ ID No. 2112	4181.1	Contig55	m	185668	186306	-	+		+
SEQ ID No. 903	2354.2	Contig55	m	186307	186954	-	-	99%	-
SEQ ID No. 543	1748.3	Contig55	m	187301	188650	-	-	96%	-

SEQ ID No. 544	1749.2	Contig55	m	188742	189518		-	91%	-
SEQ ID No. 2110	4179.2	Contig55	m	189807	191867	-	+		+
SEQ ID No. 252	1295.1	Contig55	m	191978	192574	+	-	96%	-
SEQ ID No. 251	1294.2	Contig55	m	192553	193086	-	-	98%	-
SEQ ID No. 250	1293.4	Contig55	m	193087	193707	+	-	96%	-
SEQ ID No. 2214	4337.3	Contig55	m	193708	194001	+	-	97%	-
SEQ ID No. 2215	4338.3	Contig55	m	194087	194995	-	-	97%	-
SEQ ID No. 2216	4339.1	Contig55	m	194986	196857	-	-	97%	-
SEQ ID No. 2218	4340.2	Contig55	m	196896	198794	-	-	97%	-
SEQ ID No. 2859	5386.1	Contig55	m	198790	199812	-	-	97%	-
SEQ ID No. 2532	4786.2	Contig55	m	199797	201254	+	-	98%	-
SEQ ID No. 2531	4785.1	Contig55	m	201664	202824	-	-	96%	+
SEQ ID No. 2506	4751.2	Contig55	m	202977	204695	-	-	97%	-
SEQ ID No. 387	1516.3	Contig55	m	204692	205651	-	-	96%	-
SEQ ID No. 386	1515.2	Contig55	p	205797	206564	-	-	96%	-
SEQ ID No. 385	1513.4	Contig55	p	206522	207880	+	-	98%	-
SEQ ID No. 2679	5022.2	Contig55	p	207980	209053	-	-	99%	-
SEQ ID No. 1967	3988.3	Contig55	p	209183	210673	-	-	96%	-
SEQ ID No. 1966	3986.1	Contig55	m	210767	211888	-	-	97%	-
SEQ ID No. 1965	3985.1	Contig55	p	212031	212642		-	99%	-
SEQ ID No. 3348	856.2	Contig55	p	212770	213318	-	-	97%	+
SEQ ID No. 3349	857.2	Contig55	m	213377	215725	+	-	98%	-
SEQ ID No. 2612	4908.3	Contig55	p	215762	221632	+	-	98%	-
SEQ ID No. 1974	3996.2	Contig55	m	219910	220161		-	93%	+
SEQ ID No. 494	1671.3	Contig55	m	221880	223283	+	-	96%	+
SEQ ID No. 1973	3995.5	Contig55	m	223444	224850	-	-	98%	-
SEQ ID No. 1972	3994.5	Contig55	m	224983	225843	-	-	98%	-
SEQ ID No. 1971	3993.2	Contig55	p	226063	226629	-	-	87%	+
SEQ ID No. 3372	887.4	Contig55	m	226955	229195	-	-	99%	-
SEQ ID No. 3028	5711.2	Contig55	p	229465	230373	-	-	98%	+
SEQ ID No. 2772	5193.2	Contig55	p	230456	232594	-	-	97%	-
SEQ ID No. 1938	3938.2	Contig55	m	232591	233181	-	-	98%	-
SEQ ID No. 1937	3937.2	Contig55	m	233166	234512	-	-	98%	-
SEQ ID No. 854	2272.3	Contig55	m	234482	235996	-	-	99%	-
SEQ ID No. 855	2274.2	Contig55	m	236004	237023	-	-	99%	-
SEQ ID No. 1936	3936.2	Contig55	m	236965	238269	-	-	99%	-

SEQ ID No. 684	1972.3	Contig55	m	238224	239414	-	-	99%		-
SEQ ID No. 1186	278.3	Contig55	m	239267	241183		-	98%		-
SEQ ID No. 2773	5194.1	Contig55	m	241056	241625	-	+			-
SEQ ID No. 1178	277.1	Contig55	m	241772	242491	-	-	100%		-
SEQ ID No. 1171	276.1	Contig55	m	242556	242792		-	100%		-
SEQ ID No. 547	1754.3	Contig55	m	243052	244590		-	98%		-
SEQ ID No. 548	1755.3	Contig55	m	244529	244864		-	98%		+
SEQ ID No. 549	1756.4	Contig55	m	244943	245797	-	-	98%		-
SEQ ID No. 1114	2669.2	Contig55	m	246338	247576	-	-	99%		+
SEQ ID No. 1116	2670.1	Contig55	m	247577	248836	-	-	99%		-
SEQ ID No. 1117	2671.1	Contig55	p	248990	249631	-	-	97%		-
SEQ ID No. 1118	2674.1	Contig55	p	249701	252391	-	-	99%	+	-
SEQ ID No. 1119	2677.2	Contig55	p	252653	254050	+	-	100%		-
SEQ ID No. 3344	847.4	Contig55	m	254244	257435	-	-	99%		-
SEQ ID No. 1121	2680.1	Contig55	p	257935	258630	+	-	99%		-
SEQ ID No. 518	1708.4	Contig55	p	259050	259427	-	-	99%		-
SEQ ID No. 519	1709.1	Contig55	m	259455	260177	-	-	98%		-
SEQ ID No. 520	1711.2	Contig55	m	260134	260889	-	-	99%		-
SEQ ID No. 864	2292.2	Contig55	p	261111	262802	-	-	95%		+
SEQ ID No. 1122	2682.1	Contig55	p	262935	264578	-	-	93%		+
SEQ ID No. 80	1036.4	Contig55	m	264665	266002	-	-	98%		-
SEQ ID No. 81	1037.1	Contig55	m	265921	266478	-	-	100%		-
SEQ ID No. 82	1039.2	Contig55	p	266605	267834	-	-	95%		-
SEQ ID No. 777	2144.3	Contig55	p	267780	268817	-	-	91%		-
SEQ ID No. 1123	2683.3	Contig55	p	268818	270377	-	-	96%		-
SEQ ID No. 2774	5195.2	Contig55	p	270293	271729	-	-	98%		-
SEQ ID No. 1124	2685.2	Contig55	p	271681	272466		-	98%		+
SEQ ID No. 2492	4732.3	Contig55	m	272544	272921		-	97%		-
SEQ ID No. 540	1743.4	Contig55	m	273331	273900	-	-	99%		-
SEQ ID No. 539	1742.3	Contig55	m	273878	274645	-	-	98%		-
SEQ ID No. 2493	4733.2	Contig55	m	274618	275526	-	-	98%		-
SEQ ID No. 2665	4999.2	Contig55	m	275562	276383	-	-	99%		-
SEQ ID No. 2668	5000.3	Contig55	p	276648	277070	+	-	98%		-
SEQ ID No. 2669	5001.1	Contig55	p	277565	278038	-	-	99%		+
SEQ ID No. 2820	5307.3	Contig55	m	278103	279803	-	-	97%		-
SEQ ID No. 1970	3991.3	Contig55	m	280182	281885		-	95%		-

SEQ ID No. 3336	837.2	Contig55	p	281957	282727	-	-	100%	-
SEQ ID No. 3335	835.2	Contig55	p	282705	285299	-	-	99%	-
SEQ ID No. 1968	3989.1	Contig55	m	285321	285770	-	-	99%	-
SEQ ID No. 195	1210.2	Contig55	m	285761	287344	-	-	97%	-
SEQ ID No. 193	1209.2	Contig55	m	287341	288849	-	-	100%	-
SEQ ID No. 192	1208.1	Contig55	p	288191	288577		-	99%	-
SEQ ID No. 2821	5309.1	Contig55	p	288995	289852	-	-	100%	-
SEQ ID No. 1016	2514.3	Contig55	p	289828	290118	-	-	100%	-
SEQ ID No. 3376	893.4	Contig55	m	290434	292959	+	-	99%	-
SEQ ID No. 2310	4472.2	Contig55	m	293013	293819	-	-	99%	-
SEQ ID No. 2501	4744.3	Contig55	m	293795	294718	-	-	98%	-
SEQ ID No. 378	1503.4	Contig55	m	294663	295850	-	-	97%	-
SEQ ID No. 377	1501.3	Contig55	m	295754	296599	+	-	98%	-
SEQ ID No. 928	2388.3	Contig55	m	296578	297459	-	-	97%	-
SEQ ID No. 927	2387.3	Contig55	p	297507	298577	-	-	99%	-
SEQ ID No. 154	1144.3	Contig55	m	298751	300040	-	-	99%	-
SEQ ID No. 155	1145.3	Contig55	m	300027	302186	-	-	97%	-
SEQ ID No. 2823	5311.1	Contig55	p	302427	302744	-	-	100%	-
SEQ ID No. 2044	4094.2	Contig55	p	302741	304198	-	-	97%	-
SEQ ID No. 2043	4093.1	Contig55	p	304199	305635	-	-	98%	-
SEQ ID No. 893	2343.2	Contig55	p	305819	306838	-	-	98%	-
SEQ ID No. 894	2344.4	Contig55	p	306980	309283	-	-	97%	-
SEQ ID No. 2906	5478.2	Contig55	p	309284	309784		-	88%	-
SEQ ID No. 136	1115.4	Contig55	p	309795	310592		-	97%	-
SEQ ID No. 135	1113.2	Contig55	p	310647	311045		-	96%	-
SEQ ID No. 2610	4902.1	Contig55	p	310700	311026		+		-
SEQ ID No. 2905	5477.1	Contig55	m	311164	311460	-	-	100%	-
SEQ ID No. 2904	5476.1	Contig55	m	311846	312241	-	-	98%	-
SEQ ID No. 1698	3556.2	Contig55	m	312219	312647		-	97%	-
SEQ ID No. 1699	3557.1	Contig55	m	312599	313828	-	-	98%	-
SEQ ID No. 2168	427.3	Contig55	m	313737	314516	-	-	99%	-
SEQ ID No. 2173	428.1	Contig55	p	314611	315411	-	-	100%	-
SEQ ID No. 2180	429.2	Contig55	m	315524	316543		-	98%	-
SEQ ID No. 748	2092.2	Contig55	p	316557	319145	-	-	99%	-
SEQ ID No. 747	2091.1	Contig55	m	316811	317566		-	97%	-
SEQ ID No. 1700	3558.1	Contig55	m	317987	318523		-	98%	+

SEQ ID No. 1701	3559.2	Contig55	m	319237	320562	-	-	97%	+
SEQ ID No. 635	1886.2	Contig55	m	320563	321231	-	-	99%	+
SEQ ID No. 634	1885.3	Contig55	m	321503	322891	+	-	99%	-
SEQ ID No. 2432	4649.1	Contig55	m	322872	323477		-	96%	+
SEQ ID No. 2434	4650.3	Contig55	m	323544	327368	+	-	98%	+
SEQ ID No. 2164	4262.2	Contig55	m	327355	327813	-	-	98%	-
SEQ ID No. 1039	2550.2	Contig55	m	327807	329195	-	-	98%	-
SEQ ID No. 1040	2551.2	Contig55	m	329159	329809	-	-	97%	-
SEQ ID No. 1041	2553.3	Contig55	m	329939	330940	-	-	100%	-
SEQ ID No. 2162	4259.2	Contig55	m	330931	331806		-	97%	-
SEQ ID No. 1784	3674.2	Contig55	m	331754	332569	-	+		-
SEQ ID No. 1785	3675.1	Contig55	m	332739	333086	-	-	98%	-
SEQ ID No. 1786	3676.2	Contig55	m	333068	334462	-	-	99%	-
SEQ ID No. 1787	3679.3	Contig55	m	334405	335460	-	-	99%	-
SEQ ID No. 2979	561.3	Contig55	m	335735	337042	-	-	97%	-
SEQ ID No. 2983	562.1	Contig55	m	337055	337672	+	-	99%	-
SEQ ID No. 3130	6007.1	Contig55	p	337580	337705	-	-	100%	+
SEQ ID No. 2994	564.2	Contig55	m	337666	340092	-	-	98%	-
SEQ ID No. 1789	3680.1	Contig55	p	339922	341073	+	-	98%	-
SEQ ID No. 1790	3681.1	Contig55	p	341052	341744	-	-	97%	-
SEQ ID No. 1791	3682.1	Contig55	p	341754	342179	-	-	95%	-
SEQ ID No. 1792	3683.2	Contig55	p	342284	342532	-	-	100%	-
SEQ ID No. 2817	5295.1	Contig55	m	342696	344141	-	-	99%	-
SEQ ID No. 972	2448.3	Contig55	p	344088	344876	+	-	97%	-
SEQ ID No. 971	2447.3	Contig55	p	345075	346076	-	-	98%	-
SEQ ID No. 970	2445.4	Contig55	p	346267	349176	-	-	99%	-
SEQ ID No. 2539	4797.1	Contig55	p	349121	350293	-	-	99%	-
SEQ ID No. 2538	4796.1	Contig55	p	350604	351089	+	-	96%	-
SEQ ID No. 2537	4795.1	Contig55	p	350815	351099		-	87%	+
SEQ ID No. 1983	4005.2	Contig55	m	351176	352696	-	-	99%	-
SEQ ID No. 1984	4009.1	Contig55	m	352663	353715	-	-	99%	-
SEQ ID No. 1986	4010.1	Contig55	m	353830	354234		-	100%	+
SEQ ID No. 1987	4012.1	Contig55	m	354228	355055	-	-	98%	-
SEQ ID No. 1988	4013.1	Contig55	m	355006	355788	-	-	100%	-
SEQ ID No. 1989	4014.1	Contig55	m	355789	356577	-	-	99%	-
SEQ ID No. 1990	4015.1	Contig55	m	356602	357495	-	-	99%	-

SEQ ID No. 1991	4016.1	Contig55	m	357458	358642	-	-	98%		-
SEQ ID No. 1992	4017.4	Contig55	m	358638	360716	-	-	99%		-
SEQ ID No. 2848	5361.2	Contig55	m	360713	361897	-	-	99%		-
SEQ ID No. 2847	5360.2	Contig55	p	361044	361610		-	99%		+
SEQ ID No. 1066	2592.4	Contig55	p	361844	362254		-	97%		+
SEQ ID No. 1067	2593.3	Contig55	m	361861	362634	-	-	99%		-
SEQ ID No. 2680	5026.1	Contig55	m	362635	362964	+	-	98%		-
SEQ ID No. 1068	2595.4	Contig55	m	362913	363662	+	-	97%		-
SEQ ID No. 2845	5359.2	Contig55	m	363663	364091	-	-	98%		-
SEQ ID No. 2352	4526.3	Contig55	m	364070	364414	-	-	96%		-
SEQ ID No. 434	1585.2	Contig55	m	364415	365494	-	-	95%		-
SEQ ID No. 435	1586.2	Contig55	m	365921	366190	+	-	100%		+
SEQ ID No. 436	1587.4	Contig55	p	366186	367574		-	97%		-
SEQ ID No. 2171	4275.2	Contig55	p	367445	368293	+	-	97%		-
SEQ ID No. 2170	4273.1	Contig55	p	368190	369320		-	97%		-
SEQ ID No. 876	2313.2	Contig55	m	369399	370625	-	-	96%		-
SEQ ID No. 875	2312.1	Contig55	m	370829	371116	-	-	98%		+
SEQ ID No. 874	2311.4	Contig55	m	371117	373702	-	-	98%		-
SEQ ID No. 627	1873.2	Contig55	m	373717	374169	-	-	100%		-
SEQ ID No. 626	1871.2	Contig55	m	374162	375211	-	-	100%		-
SEQ ID No. 2428	4643.2	Contig55	m	375511	376461	-	-	99%		+
SEQ ID No. 2427	4642.1	Contig55	m	376740	377243	-	-	98%	+	-
SEQ ID No. 3101	5910.1	Contig55	m	377311	377523		-	98%		+
SEQ ID No. 2426	4640.2	Contig55	p	377831	380395	-	-	99%		-
SEQ ID No. 1994	4020.2	Contig55	p	380396	382096		-	99%		-
SEQ ID No. 3332	830.4	Contig55	p	382039	384633		-	95%		+
SEQ ID No. 3333	832.3	Contig55	p	384667	385662	-	-	94%		-
SEQ ID No. 3041	5747.1	Contig55	p	385787	386191	+	-	96%		+
SEQ ID No. 678	1955.3	Contig55	m	386532	388040	+	-	92%		-
SEQ ID No. 885	2328.2	Contig55	p	388278	390869	-	-	99%		+
SEQ ID No. 1501	3269.1	Contig55	p	390859	394104	-	-	98%		-
SEQ ID No. 960	2430.2	Contig55	p	394309	395385	-	-	99%		-
SEQ ID No. 1500	3267.1	Contig55	p	395435	396055		-	99%		+
SEQ ID No. 1499	3265.1	Contig55	m	396103	396999	-	-	98%		-
SEQ ID No. 1498	3263.2	Contig55	p	397202	398425	-	-	98%		-
SEQ ID No. 1497	3260.2	Contig55	m	398427	399434	+	-	99%		+

SEQ ID No. 467	1633.2	Contig55	m	399379	400365	-	-	99%	-	
SEQ ID No. 466	1632.3	Contig55	m	400350	402146	+	-	99%	-	
SEQ ID No. 2811	5278.3	Contig55	m	402203	403054	-	-	99%	+	
SEQ ID No. 2810	5277.3	Contig55	m	403006	404235	-	-	100%	-	
SEQ ID No. 1371	3070.2	Contig55	m	404293	405312	+	-	100%	+	
SEQ ID No. 1369	3069.1	Contig55	m	405340	406062	-	-	99%	-	
SEQ ID No. 1368	3068.1	Contig55	p	406331	407521	-	-	100%	-	
SEQ ID No. 1031	2538.2	Contig55	p	407509	408717	-	-	100%	-	
SEQ ID No. 1367	3067.1	Contig55	p	408846	409169		-	100%	-	
SEQ ID No. 3015	568.2	Contig55	p	409176	410786	-	-	99%	-	
SEQ ID No. 3020	569.2	Contig55	p	410743	411600	-	-	99%	-	
SEQ ID No. 3027	571.2	Contig55	p	411503	413557	-	-	97%	-	
SEQ ID No. 699	2007.1	Contig55	p	413535	414476	-	-	100%	-	
SEQ ID No. 145	1127.3	Contig55	p	414552	416573	-	-	98%	-	
SEQ ID No. 1366	3065.1	Contig55	m	416733	417197		-	96%	-	
SEQ ID No. 1365	3063.2	Contig55	p	417443	418627	-	-	99%	-	
SEQ ID No. 1364	3061.2	Contig55	p	418602	419252	-	-	99%	-	
SEQ ID No. 1363	3060.1	Contig55	m	418655	419038		-	96%	+	
SEQ ID No. 1361	3059.2	Contig55	p	419253	420044	+	-	99%	-	
SEQ ID No. 95	1058.3	Contig55	p	420079	421542	-	-	98%	+	
SEQ ID No. 94	1056.1	Contig55	p	421787	422035	-	-	100%	+	
SEQ ID No. 93	1055.1	Contig55	m	422032	422802	-	-	97%	-	
SEQ ID No. 92	1053.2	Contig55	m	422766	423317	-	-	99%	-	
SEQ ID No. 521	1713.2	Contig55	m	423302	425392	-	-	97%	-	
SEQ ID No. 522	1714.4	Contig55	m	425358	426290	-	-	100%	-	
SEQ ID No. 579	1797.4	Contig55	m	426498	427307	+	-	100%	-	
SEQ ID No. 582	1800.4	Contig55	m	427491	429701	-	-	98%	-	
SEQ ID No. 2277	4419.2	Contig55	p	429883	430851	-	-	99%	-	
SEQ ID No. 2899	5469.1	Contig55	p	430842	431012		+		-	
SEQ ID No. 2900	5470.1	Contig55	p	430963	431283	-	-	100%	-	
SEQ ID No. 228	1260.3	Contig55	p	431468	432283	+	-	100%	-	
SEQ ID No. 226	1258.1	Contig55	m	432392	433441	-	-	97%	-	
SEQ ID No. 225	1257.5	Contig55	m	433351	434646	-	-	99%	-	
SEQ ID No. 3129	6004.1	Contig55	m	434692	434850		+		+	
SEQ ID No. 2901	5471.2	Contig55	m	434805	435017	-	+		-	
SEQ ID No. 3099	5907.1	Contig55	m	434953	435111		+		-	

SEQ ID No. 3040	5743.1	Contig55	m	435125	435313		+			+
SEQ ID No. 2895	5458.1	Contig56	m	1	294		+			-
SEQ ID No. 1829	3744.2	Contig56	p	629	1834		+			+
SEQ ID No. 1828	3743.1	Contig56	m	1999	2334	+	-	96%		+
SEQ ID No. 1827	3742.1	Contig56	m	2352	2624		-	92%		+
SEQ ID No. 1826	3741.1	Contig56	m	2635	3855	-	-	98%		+
SEQ ID No. 1825	3740.3	Contig56	m	4089	4625	-	-	100%		-
SEQ ID No. 1823	3739.3	Contig56	p	4848	5291	-	-	99%		-
SEQ ID No. 1822	3737.1	Contig56	p	5344	6447	+	-	94%		+
SEQ ID No. 113	1082.3	Contig56	m	6468	8033	-	+			-
SEQ ID No. 357	1469.4	Contig56	m	8879	12292	-	+			+
SEQ ID No. 2197	4312.2	Contig56	m	12373	13725	-	-	98%		-
SEQ ID No. 2196	4311.1	Contig56	p	13958	14506	-	-	96%		-
SEQ ID No. 2195	4310.1	Contig56	m	14610	15161	-	-	97%		+
SEQ ID No. 2193	4309.2	Contig56	m	15077	16219		-	97%		+
SEQ ID No. 2010	4047.3	Contig56	m	16291	17385	+	-	99%		-
SEQ ID No. 2011	4048.3	Contig56	p	17493	19334	-	-	100%		-
SEQ ID No. 2013	4050.1	Contig56	p	19403	20197	+	-	98%		-
SEQ ID No. 2014	4051.2	Contig56	p	20206	20601	+	-	98%		-
SEQ ID No. 2015	4054.2	Contig56	p	20585	21265	-	-	99%		-
SEQ ID No. 2016	4055.1	Contig56	m	21293	23185	-	+		+	-
SEQ ID No. 2017	4056.2	Contig56	m	23266	24495	-	-	99%		-
SEQ ID No. 2018	4058.2	Contig56	m	24583	24942		+			+
SEQ ID No. 472	1641.5	Contig56	p	25345	26694	-	-	98%		-
SEQ ID No. 473	1642.3	Contig56	p	26679	27341	-	-	100%		-
SEQ ID No. 2465	4699.3	Contig56	p	27452	28756	-	-	99%		-
SEQ ID No. 2340	4510.1	Contig56	p	28978	31440	-	-	100%		-
SEQ ID No. 2341	4511.1	Contig56	p	31527	31847	-	-	98%	+	-
SEQ ID No. 2342	4512.4	Contig56	p	32084	33979	-	-	98%		-
SEQ ID No. 572	1787.4	Contig56	p	34519	37686	-	+			+
SEQ ID No. 1678	3530.1	Contig56	m	37969	38598	-	+			+
SEQ ID No. 1679	3532.1	Contig56	p	38652	41942	-	+			+
SEQ ID No. 368	1487.3	Contig56	m	42394	43230	-	-	99%	+	-
SEQ ID No. 1680	3533.1	Contig56	m	43296	44021	+	-	99%		-
SEQ ID No. 1681	3534.2	Contig56	m	43987	45666	+	-	99%		-
SEQ ID No. 1682	3535.5	Contig56	p	45982	46656	-	-	95%		+

SEQ ID No. 1683	3536.2	Contig56	m	46690	47052	-	-	98%		-
SEQ ID No. 2732	5098.2	Contig56	m	47027	47311	-	-	100%		+
SEQ ID No. 2731	5097.1	Contig56	m	47312	48241	+	-	99%		-
SEQ ID No. 3128	6002.1	Contig56	p	48661	49746	-	+			-
SEQ ID No. 3125	5999.1	Contig56	p	49760	50065	-	-	100%		-
SEQ ID No. 2480	4715.3	Contig56	p	50326	50817		-	95%		-
SEQ ID No. 2479	4714.1	Contig56	m	50839	51837	-	-	96%		-
SEQ ID No. 2478	4713.1	Contig56	p	51969	52322	-	-	93%		+
SEQ ID No. 2935	5532.2	Contig56	m	52635	53546	-	-	97%		-
SEQ ID No. 2936	5533.2	Contig56	m	53673	53912		+			-
SEQ ID No. 499	1682.4	Contig56	p	54221	55105		-	94%		-
SEQ ID No. 500	1683.2	Contig56	m	55230	55457		+			+
SEQ ID No. 501	1684.3	Contig56	p	55649	56050	-	-	96%		+
SEQ ID No. 2249	4378.2	Contig56	p	56210	56911	-	-	95%		-
SEQ ID No. 2250	4379.1	Contig56	p	56884	57786	-	-	94%		-
SEQ ID No. 2251	4381.2	Contig56	p	57812	59134	+	-	99%		-
SEQ ID No. 2252	4383.2	Contig56	p	59183	60082	+	-	97%		-
SEQ ID No. 997	2483.2	Contig56	m	60143	60406		-	93%		+
SEQ ID No. 996	2482.1	Contig56	m	60536	61048		-	96%		-
SEQ ID No. 995	2481.4	Contig56	m	61231	62142	-	-	97%		-
SEQ ID No. 2401	4600.2	Contig56	m	62433	63200	-	-	97%		-
SEQ ID No. 2402	4601.1	Contig56	m	63160	63900	-	-	97%		+
SEQ ID No. 2403	4602.3	Contig56	p	63958	64875	-	-	98%		-
SEQ ID No. 3002	5652.2	Contig56	p	64957	65739	-	-	97%		+
SEQ ID No. 2009	4045.2	Contig56	p	65724	66335	-	+			+
SEQ ID No. 2008	4043.1	Contig56	m	66641	67078	-	-	98%		-
SEQ ID No. 2007	4042.1	Contig56	p	67106	67858	-	-	96%		-
SEQ ID No. 2006	4041.1	Contig56	p	68024	69148	+	-	96%		+
SEQ ID No. 2005	4040.1	Contig56	p	69187	69486		+			-
SEQ ID No. 2004	4039.2	Contig56	m	69539	70852	-	+			-
SEQ ID No. 2003	4037.2	Contig56	m	70837	71115	-	+			-
SEQ ID No. 2002	4036.2	Contig56	p	71116	71400		+			-
SEQ ID No. 2001	4035.1	Contig56	p	71831	72199	-	-	80%	+	-
SEQ ID No. 2326	4490.2	Contig56	m	72497	73351	-	+			-
SEQ ID No. 2327	4491.3	Contig56	m	73671	73943	-	+			+
SEQ ID No. 2328	4492.3	Contig56	p	74457	75404		+			+

SEQ ID No. 2329	4493.1	Contig56	p	75398	75607	-	+			+
SEQ ID No. 967	2441.2	Contig56	p	75770	77053	-	+			-
SEQ ID No. 965	2439.2	Contig56	p	77137	77460	-	+			-
SEQ ID No. 964	2438.2	Contig56	p	77301	77741	-	+			-
SEQ ID No. 2330	4496.2	Contig56	p	77914	78756	-	+			+
SEQ ID No. 2975	5605.2	Contig56	p	79041	79460	-	+			+
SEQ ID No. 425	1570.4	Contig56	p	79550	82762	+	+		+	-
SEQ ID No. 2343	4516.3	Contig56	p	82740	83900	+	+			-
SEQ ID No. 306	1387.2	Contig56	p	83888	85336	-	+			-
SEQ ID No. 307	1388.2	Contig56	m	85597	85953		+			+
SEQ ID No. 853	2271.1	Contig56	m	85954	86205		+			+
SEQ ID No. 852	2270.2	Contig56	m	86198	86578		+			-
SEQ ID No. 2860	5387.1	Contig56	p	86587	87405	-	+			-
SEQ ID No. 2861	5388.1	Contig56	p	87345	87626	-	+			+
SEQ ID No. 1941	3953.1	Contig56	m	87959	89164	-	-	99%		-
SEQ ID No. 506	1692.2	Contig56	p	89181	91175	-	-	98%	+	-
SEQ ID No. 1940	3950.2	Contig56	p	91527	93143	+	-	94%		-
SEQ ID No. 2681	5028.2	Contig56	m	93235	94818	+	-	97%	+	-
SEQ ID No. 2863	5390.1	Contig56	m	94884	95249		-	97%		-
SEQ ID No. 1995	4021.2	Contig56	p	95342	95935	-	-	98%		-
SEQ ID No. 1996	4023.2	Contig56	p	95838	97349	-	-	99%		-
SEQ ID No. 1997	4025.1	Contig56	p	97336	98541	-	-	98%		-
SEQ ID No. 1998	4026.2	Contig56	p	98511	99701	+	-	97%		-
SEQ ID No. 273	1334.3	Contig56	p	99838	100644	-	-	86%		+
SEQ ID No. 274	1335.2	Contig56	m	100634	101995		-	90%		+
SEQ ID No. 1999	4030.1	Contig56	p	102167	102703	-	-	79%		+
SEQ ID No. 2000	4032.3	Contig56	m	102700	103668	-	-	99%		-
SEQ ID No. 2864	5391.2	Contig56	m	103725	105176	-	-	99%		-
SEQ ID No. 3124	5996.1	Contig56	m	105163	105690	+	-	98%		-
SEQ ID No. 2822	531.4	Contig56	p	105919	107865	-	-	98%		+
SEQ ID No. 2819	530.3	Contig56	m	107831	108436	-	-	96%		-
SEQ ID No. 285	1356.2	Contig56	m	108648	109430		-	88%		-
SEQ ID No. 286	1357.2	Contig56	m	109467	109916	-	-	98%		-
SEQ ID No. 2635	4948.4	Contig56	p	110012	113998	-	-	96%		+
SEQ ID No. 869	2299.3	Contig56	p	114216	114683	-	-	95%		-
SEQ ID No. 870	2301.2	Contig56	p	115076	117259	-	-	99%		-

SEQ ID No. 1354	3049.1	Contig56	p	117222	118001	-	-	97%		-
SEQ ID No. 1353	3047.2	Contig56	p	117973	118644	-	-	99%		-
SEQ ID No. 1352	3046.3	Contig56	m	118221	118781		-	97%		-
SEQ ID No. 1351	3045.3	Contig56	p	118645	120087	-	-	99%		-
SEQ ID No. 1350	3042.1	Contig56	m	120097	120294	-	+			+
SEQ ID No. 1349	3041.1	Contig56	p	120462	121664	-	-	95%		+
SEQ ID No. 2967	559.2	Contig56	p	121830	122159		-	97%		-
SEQ ID No. 2961	558.2	Contig56	p	122262	122726	-	-	98%		-
SEQ ID No. 2959	557.2	Contig56	m	122756	125467	-	-	98%		-
SEQ ID No. 606	1844.3	Contig56	m	125449	126537	-	-	100%		-
SEQ ID No. 607	1845.3	Contig56	p	126753	127550	+	-	100%		+
SEQ ID No. 608	1846.3	Contig56	m	127547	128779	-	-	99%		-
SEQ ID No. 71	1022.2	Contig56	m	128901	129779	-	-	97%		-
SEQ ID No. 72	1023.2	Contig56	m	129880	130467	+	-	95%		-
SEQ ID No. 73	1024.2	Contig56	m	130563	131165	-	-	99%		-
SEQ ID No. 348	1451.4	Contig56	m	131482	132765	-	-	98%		-
SEQ ID No. 617	1859.2	Contig56	m	132780	133634	-	-	100%		-
SEQ ID No. 616	1858.1	Contig56	m	133685	133957	-	-	100%		+
SEQ ID No. 615	1856.3	Contig56	p	134558	136363	-	-	99%		+
SEQ ID No. 2383	4574.2	Contig56	p	136288	137400		-	95%		-
SEQ ID No. 164	1160.2	Contig56	m	137686	138921	-	+			-
SEQ ID No. 162	1159.5	Contig56	m	139130	140416	-	-	99%		+
SEQ ID No. 498	1681.5	Contig56	p	141000	143876	-	-	98%		-
SEQ ID No. 994	248.2	Contig56	m	143990	145393	-	-	100%	+	-
SEQ ID No. 987	247.1	Contig56	m	145394	146179	-	-	99%		-
SEQ ID No. 973	245.1	Contig56	m	146315	147349	-	-	100%	+	-
SEQ ID No. 966	244.1	Contig56	m	147447	148508	-	-	99%		-
SEQ ID No. 959	243.2	Contig56	m	148517	149470	-	-	99%		-
SEQ ID No. 702	2017.1	Contig56	p	149611	150183	-	-	98%		-
SEQ ID No. 701	2016.2	Contig56	p	150223	150675	-	-	100%		-
SEQ ID No. 1343	3028.1	Contig56	p	150665	150934	-	-	100%		-
SEQ ID No. 1342	3027.1	Contig56	p	150941	151432	-	-	99%		-
SEQ ID No. 1837	376.3	Contig56	p	151338	152423	+	-	98%		-
SEQ ID No. 1843	377.1	Contig56	m	152501	153340	-	-	99%		-
SEQ ID No. 1865	381.6	Contig56	m	153410	156961	-	-	89%		+
SEQ ID No. 3090	5899.1	Contig56	p	157044	157373		-	100%		+

SEQ ID No. 595	1823.5	Contig56	m	157377	158678	-	-	99%	+	-
SEQ ID No. 2223	4346.2	Contig56	m	158924	159877	-	-	99%		-
SEQ ID No. 2224	4347.2	Contig56	p	159936	160262	+	-	98%		-
SEQ ID No. 2225	4349.2	Contig56	p	160489	161322	+	-	100%		-
SEQ ID No. 2226	4350.2	Contig56	p	161323	162489	-	-	98%		-
SEQ ID No. 2227	4351.2	Contig56	p	162510	163424	-	-	98%		-
SEQ ID No. 2228	4352.3	Contig56	p	163425	164372	-	-	98%		-
SEQ ID No. 3121	5986.2	Contig56	m	164438	164929	-	-	98%		+
SEQ ID No. 2460	4691.2	Contig56	p	165075	166451	-	-	99%		-
SEQ ID No. 2459	4690.2	Contig56	m	166544	168484	-	-	98%		+
SEQ ID No. 2457	4688.2	Contig56	m	168556	169680	-	-	98%	+	-
SEQ ID No. 2012	405.3	Contig56	m	169767	171719	-	-	98%		-
SEQ ID No. 2019	406.1	Contig56	m	171982	172845	-	-	98%		-
SEQ ID No. 2027	407.4	Contig56	m	172815	174182	-	-	99%	+	-
SEQ ID No. 2874	5420.1	Contig56	p	174276	174809	-	+			-
SEQ ID No. 2875	5421.1	Contig56	p	174790	175281		-	97%		+
SEQ ID No. 2876	5422.1	Contig56	p	175516	175929	-	-	96%		+
SEQ ID No. 2877	5423.1	Contig56	p	176093	176353	-	+			-
SEQ ID No. 825	2225.2	Contig56	p	176301	177230	-	-	97%		-
SEQ ID No. 146	1128.3	Contig56	m	177274	178308	-	-	99%		-
SEQ ID No. 147	1132.2	Contig56	p	178079	178876	-	-	97%		-
SEQ ID No. 148	1133.3	Contig56	m	178905	180047	-	-	97%		-
SEQ ID No. 149	1134.3	Contig56	m	180035	180313	-	-	100%		+
SEQ ID No. 1861	3802.2	Contig56	p	180508	181905	-	-	100%		-
SEQ ID No. 1862	3803.1	Contig56	p	181817	182977	-	-	100%		-
SEQ ID No. 1863	3804.2	Contig56	p	182964	183833	-	-	98%		-
SEQ ID No. 1864	3807.2	Contig56	p	183917	184408	-	-	99%		-
SEQ ID No. 321	1407.2	Contig56	p	184433	185311	-	-	98%		+
SEQ ID No. 320	1406.2	Contig56	m	185316	186611	-	-	96%		-
SEQ ID No. 319	1405.2	Contig56	p	186721	187620	-	-	98%		-
SEQ ID No. 2878	5424.2	Contig56	m	187587	188366		-	98%		+
SEQ ID No. 2749	5134.4	Contig56	m	188552	189529	+	-	99%		-
SEQ ID No. 2748	5133.3	Contig56	p	189601	190266	-	-	99%		+
SEQ ID No. 2747	5132.3	Contig56	m	190423	190665	-	+			+
SEQ ID No. 1372	3072.2	Contig56	m	190992	192803	-	-	99%		-
SEQ ID No. 3203	640.2	Contig56	m	193129	194091	-	-	97%		-

SEQ ID No. 3204	641.3	Contig56	m	194043	196091	-	-	99%	-
SEQ ID No. 457	1618.3	Contig56	p	196354	196860	+	-	98%	+
SEQ ID No. 456	1617.3	Contig56	m	196904	197545	-	-	98%	-
SEQ ID No. 785	2152.2	Contig56	m	197442	198038	-	-	97%	-
SEQ ID No. 1373	3074.1	Contig56	p	198106	199032	-	-	96%	+
SEQ ID No. 631	1880.3	Contig56	m	199130	200137	-	-	100%	-
SEQ ID No. 979	2459.1	Contig56	m	200112	200678	-	-	99%	-
SEQ ID No. 978	2458.2	Contig56	p	200733	201881	-	-	98%	-
SEQ ID No. 791	2162.2	Contig56	m	201913	203340	-	-	98%	-
SEQ ID No. 792	2164.2	Contig56	p	203649	204509	-	-	98%	-
SEQ ID No. 1374	3075.2	Contig56	m	204200	204562		-	95%	+
SEQ ID No. 646	1902.5	Contig56	m	204808	205812	-	-	98%	-
SEQ ID No. 2849	5369.2	Contig56	p	206011	206253	-	-	100%	-
SEQ ID No. 2851	5370.1	Contig56	p	206447	206899	-	-	98%	-
SEQ ID No. 860	2282.4	Contig56	p	206833	208641	-	-	99%	-
SEQ ID No. 169	1169.3	Contig56	p	208701	210593	-	-	99%	-
SEQ ID No. 168	1168.2	Contig56	m	209799	210206		-	94%	-
SEQ ID No. 3077	5871.3	Contig56	m	210934	212142	-	+		-
SEQ ID No. 1824	374.2	Contig56	m	212175	213962	-	-	90%	-
SEQ ID No. 1809	371.1	Contig56	p	214213	214563	-	+		+
SEQ ID No. 565	1779.3	Contig56	m	214773	214958		+		+
SEQ ID No. 567	1780.3	Contig56	p	215260	215511	-	-	100%	-
SEQ ID No. 568	1781.4	Contig56	p	215742	217157	-	-	98%	-
SEQ ID No. 441	1592.3	Contig56	p	217296	218225	+	-	99%	-
SEQ ID No. 2177	4285.1	Contig56	m	218497	218913		+		+
SEQ ID No. 2176	4284.2	Contig56	m	218914	219987	-	+		-
SEQ ID No. 2175	4282.2	Contig56	m	220040	220858	-	+		-
SEQ ID No. 2174	4281.2	Contig56	p	220863	221168	+	+		-
SEQ ID No. 2172	4276.2	Contig56	m	221531	223492	-	+		-
SEQ ID No. 3078	5872.1	Contig56	m	223532	223738	-	+		+
SEQ ID No. 2357	4532.2	Contig56	p	223941	225431	-	+		+
SEQ ID No. 2356	4531.2	Contig56	m	225684	226010	-	+		-
SEQ ID No. 2355	4530.1	Contig56	m	226119	226424	-	+		-
SEQ ID No. 2353	4528.2	Contig56	m	226438	227223	-	+		-
SEQ ID No. 2921	5504.4	Contig56	m	227405	228223		+		+
SEQ ID No. 906	2357.4	Contig56	p	228590	230416	-	+		-

SEQ ID No. 3355	864.1	Contig56	m	230660	231064	-	-	77%	-
SEQ ID No. 3354	863.1	Contig56	m	231159	231653	-	+		-
SEQ ID No. 3353	862.1	Contig56	p	231881	232144	-	+		+
SEQ ID No. 3352	860.2	Contig56	p	232289	234031	-	+		-
SEQ ID No. 1598	3411.2	Contig56	p	234499	235608	-	+		-
SEQ ID No. 1599	3413.2	Contig56	p	235844	236812	-	+		-
SEQ ID No. 1600	3414.3	Contig56	m	237027	238826	-	+		-
SEQ ID No. 652	1910.6	Contig56	m	239000	243505	-	+		+
SEQ ID No. 653	1911.4	Contig56	m	243805	244335	-	-	86%	-
SEQ ID No. 403	1536.3	Contig56	m	244316	246199	+	-	95%	-
SEQ ID No. 450	1603.3	Contig56	p	246579	248123	-	+		-
SEQ ID No. 1902	3878.1	Contig56	m	248247	251330	-	+		+
SEQ ID No. 1901	3876.1	Contig56	p	251633	252130		-	92%	-
SEQ ID No. 1900	3874.1	Contig56	m	252199	252504	-	-	95%	+
SEQ ID No. 1899	3872.1	Contig56	p	252637	252915		+		+
SEQ ID No. 1898	3871.1	Contig56	p	252925	253329	-	+		+
SEQ ID No. 2870	5406.1	Contig56	m	253581	255119	-	-	97%	-
SEQ ID No. 259	1307.4	Contig56	p	255297	256598	-	-	93%	-
SEQ ID No. 258	1306.5	Contig56	m	256705	259905	+	-	96%	-
SEQ ID No. 2869	5405.2	Contig56	m	259906	260877	+	-	97%	-
SEQ ID No. 59	1001.3	Contig56	m	260871	262217	+	-	95%	-
SEQ ID No. 60	1002.2	Contig56	m	262543	264180	-	-	80%	-
SEQ ID No. 2500	4742.3	Contig56	p	264520	265437	-	-	97%	-
SEQ ID No. 2510	4757.2	Contig56	m	265449	266069	-	-	92%	-
SEQ ID No. 2511	4758.2	Contig56	m	266170	266652	-	-	95%	-
SEQ ID No. 2512	4759.2	Contig56	p	266570	267409	-	-	98%	-
SEQ ID No. 2513	4760.2	Contig56	m	267786	268739	-	-	95%	+
SEQ ID No. 2709	5065.2	Contig56	p	268942	269424	-	-	96%	-
SEQ ID No. 2710	5066.2	Contig56	m	269508	270020	-	-	97%	-
SEQ ID No. 474	1643.4	Contig56	m	269917	272451	-	-	96%	-
SEQ ID No. 2335	4500.1	Contig56	p	272747	273112	-	+		-
SEQ ID No. 2332	4498.1	Contig56	p	273538	275025	-	-	97%	+
SEQ ID No. 2331	4497.4	Contig56	p	275092	276381	-	-	98%	+
SEQ ID No. 667	1937.4	Contig56	p	276466	276828	-	-	98%	+
SEQ ID No. 668	1939.5	Contig56	p	277056	277634	-	-	95%	+
SEQ ID No. 2942	5542.3	Contig56	p	277693	278568	-	-	98%	-

SEQ ID No. 1918	3904.3	Contig56	p	278644	280497	-	-	97%	-
SEQ ID No. 2267	440.4	Contig56	p	280671	285386	-	-	95%	-
SEQ ID No. 1919	3908.3	Contig56	p	285579	286721	-	-	97%	+
SEQ ID No. 1813	3714.3	Contig56	p	286765	289398	-	-	97%	+
SEQ ID No. 3324	817.7	Contig56	p	289615	295440	-	-	93%	+
SEQ ID No. 664	1933.4	Contig56	p	295715	300448	-	-	87%	+
SEQ ID No. 1328	3003.1	Contig56	p	300856	301554	-	-	99%	-
SEQ ID No. 1327	3002.1	Contig56	m	301583	302824		-	96%	-
SEQ ID No. 963	2436.2	Contig56	m	302990	304480	-	-	96%	+
SEQ ID No. 1326	3001.1	Contig56	p	304644	304823	-	+		+
SEQ ID No. 1325	3000.1	Contig56	p	304917	307223	-	-	98%	-
SEQ ID No. 1321	2999.3	Contig56	p	307366	308247	-	-	92%	-
SEQ ID No. 1320	2998.1	Contig56	m	308384	308848	+	-	97%	-
SEQ ID No. 1319	2996.1	Contig56	p	309173	309691	+	-	98%	+
SEQ ID No. 1318	2995.1	Contig56	p	309823	311085	-	-	97%	+
SEQ ID No. 656	1918.3	Contig56	m	311244	311996	-	-	97%	-
SEQ ID No. 1026	2528.3	Contig56	m	311932	312699	-	-	88%	+
SEQ ID No. 1317	2994.2	Contig56	p	312863	313255	-	-	95%	-
SEQ ID No. 1316	2993.2	Contig56	p	313301	314026		-	95%	-
SEQ ID No. 1315	2992.2	Contig56	p	314243	315259	-	-	92%	-
SEQ ID No. 1314	2991.1	Contig56	m	315382	316488	-	-	96%	-
SEQ ID No. 1313	2990.1	Contig56	p	316851	317330	-	+		-
SEQ ID No. 1312	2988.1	Contig56	p	317318	317605	-	+		+
SEQ ID No. 1311	2987.1	Contig56	m	317642	318247		+		+
SEQ ID No. 1310	2986.1	Contig56	m	318269	318838		+		+
SEQ ID No. 1309	2985.1	Contig56	m	318799	319164	-	+		+
SEQ ID No. 1308	2983.1	Contig56	p	319420	319680	-	+		-
SEQ ID No. 1307	2982.1	Contig56	m	319677	321527	-	+		-
SEQ ID No. 1306	2978.1	Contig56	m	321511	322419	-	+		-
SEQ ID No. 1305	2976.1	Contig56	m	322361	322648	-	+		+
SEQ ID No. 1304	2975.2	Contig56	m	322563	322895	-	+		+
SEQ ID No. 1303	2973.2	Contig56	m	322880	323161	-	+		+
SEQ ID No. 1302	2972.3	Contig56	m	323201	324955	-	+		-
SEQ ID No. 3013	5675.2	Contig56	m	324956	326062	-	+		-
SEQ ID No. 3014	5677.2	Contig56	m	326486	327133	-	+		-
SEQ ID No. 2796	5242.3	Contig56	m	327199	327939		+		+

SEQ ID No. 2797	5243.2	Contig56	m	327951	328175		+			+
SEQ ID No. 344	1446.3	Contig56	p	328545	329315	-	-	91%		-
SEQ ID No. 343	1444.3	Contig56	m	329451	331304	-	-	95%		-
SEQ ID No. 625	1870.3	Contig56	p	331500	332612	+	-	97%		-
SEQ ID No. 3195	630.4	Contig56	p	332655	335864	-	-	98%		-
SEQ ID No. 1460	3205.3	Contig56	p	335981	340345		-	97%		-
SEQ ID No. 1461	3207.5	Contig56	p	340451	342949	-	-	97%		-
SEQ ID No. 2563	4830.2	Contig56	p	343121	344029	-	-	99%		-
SEQ ID No. 2564	4833.2	Contig56	p	344290	345072	-	-	98%		-
SEQ ID No. 2798	5247.1	Contig56	p	345073	346131	-	-	97%		-
SEQ ID No. 190	1205.5	Contig56	p	346389	357734	-	-	97%		-
SEQ ID No. 191	1207.2	Contig56	m	357737	358141		-	99%		+
SEQ ID No. 1280	2949.2	Contig56	m	358365	359324	-	-	99%		-
SEQ ID No. 1282	2951.2	Contig56	p	359726	361783	-	+			+
SEQ ID No. 1283	2952.1	Contig56	m	361896	363770	-	-	85%		+
SEQ ID No. 1284	2953.2	Contig56	p	364225	364773	-	-	96%		-
SEQ ID No. 813	2204.2	Contig56	p	364871	366475	+	-	96%		-
SEQ ID No. 814	2207.2	Contig56	p	366476	367105	-	-	96%		-
SEQ ID No. 815	2208.1	Contig56	m	367223	367516		-	77%		+
SEQ ID No. 816	2209.4	Contig56	m	367547	368536	-	-	96%		-
SEQ ID No. 1043	2557.4	Contig56	m	368573	369052	-	-	96%		-
SEQ ID No. 1042	2555.5	Contig56	p	369375	370547	-	+			+
SEQ ID No. 2408	4613.3	Contig56	p	370702	372222	+	-	95%		+
SEQ ID No. 2407	4609.2	Contig56	m	372227	372832	-	-	87%		+
SEQ ID No. 2406	4608.1	Contig56	m	373025	373588	-	-	91%		-
SEQ ID No. 2405	4607.2	Contig56	m	373694	374998	-	-	99%	+	-
SEQ ID No. 2981	5611.1	Contig56	m	374824	375147		+			-
SEQ ID No. 2980	5610.1	Contig56	m	375175	375633	-	-	92%		-
SEQ ID No. 2978	5609.1	Contig56	m	375588	375899	-	+			-
SEQ ID No. 2723	5084.4	Contig56	m	375956	377518	-	-	96%		-
SEQ ID No. 2977	5607.2	Contig56	m	377512	378960	-	-	99%		-
SEQ ID No. 2976	5606.2	Contig56	p	379119	379787		-	91%		+
SEQ ID No. 262	1317.3	Contig56	p	379963	381123		-	98%		+
SEQ ID No. 263	1318.4	Contig56	m	381171	381926		-	95%		-
SEQ ID No. 3007	5662.1	Contig56	m	381808	382452	-	+			-
SEQ ID No. 2366	4547.4	Contig56	m	382744	383787	-	-	99%		-

SEQ ID No. 2104	4171.4	Contig56	p	384054	384602	+	-	98%		-
SEQ ID No. 2105	4172.2	Contig56	m	384751	385848	-	-	97%		-
SEQ ID No. 2106	4173.2	Contig56	m	385948	386391	-	-	98%		-
SEQ ID No. 1025	2526.3	Contig56	p	386531	387826	-	-	97%		-
SEQ ID No. 2107	4174.1	Contig56	p	388061	388519	+	-	100%		+
SEQ ID No. 2108	4175.1	Contig56	m	388616	389545	-	-	98%		-
SEQ ID No. 2109	4177.2	Contig56	m	389619	391223	-	-	97%		-
SEQ ID No. 2952	5557.2	Contig56	m	391235	393079	-	-	95%		+
SEQ ID No. 2424	4636.4	Contig56	p	393601	398376	-	+			-
SEQ ID No. 647	1903.4	Contig56	m	398612	399586	-	-	97%		-
SEQ ID No. 3051	5796.3	Contig56	p	399743	400387	-	-	96%		-
SEQ ID No. 1529	3308.4	Contig56	p	400486	400734	-	+			+
SEQ ID No. 1528	3307.4	Contig56	p	400664	400867	-	+			+
SEQ ID No. 2913	549.5	Contig56	p	400978	402165		-	99%		+
SEQ ID No. 2908	548.3	Contig56	m	402228	403745	+	-	97%		+
SEQ ID No. 1527	3306.4	Contig56	m	403870	404619	+	-	100%		-
SEQ ID No. 1526	3304.2	Contig56	m	404967	406145	+	-	99%		-
SEQ ID No. 1525	3303.2	Contig56	m	406350	407627	-	-	95%		+
SEQ ID No. 115	1088.3	Contig56	m	407795	409327	-	-	96%		-
SEQ ID No. 114	1087.2	Contig56	m	409331	411073	-	-	96%		-
SEQ ID No. 1524	3301.2	Contig56	m	411031	412473	-	-	98%		-
SEQ ID No. 1521	3299.3	Contig56	m	412674	413180	-	+			-
SEQ ID No. 3110	5950.1	Contig56	m	413084	413554		+			+
SEQ ID No. 1520	3297.3	Contig56	p	414188	415258	-	-	98%		-
SEQ ID No. 325	1412.5	Contig56	p	415252	417015	-	-	98%		-
SEQ ID No. 326	1413.2	Contig56	p	416981	418399	-	-	95%		-
SEQ ID No. 327	1414.1	Contig56	p	418371	418622	-	+			-
SEQ ID No. 328	1415.4	Contig56	p	418607	419371	-	-	97%	+	-
SEQ ID No. 1250	2900.2	Contig56	p	419278	420378	-	-	97%	+	-
SEQ ID No. 1247	2898.4	Contig56	m	420698	420934	-	+			-
SEQ ID No. 604	1840.3	Contig56	m	420935	422323	-	-	99%		-
SEQ ID No. 602	1839.2	Contig56	p	422240	423514	-	-	97%		-
SEQ ID No. 1246	2893.2	Contig56	p	423595	426009	-	-	97%		-
SEQ ID No. 1245	2891.2	Contig56	m	426200	428203	-	-	98%		-
SEQ ID No. 1244	2889.2	Contig56	m	428194	428649	+	-	99%		-
SEQ ID No. 1034	2544.3	Contig56	p	428750	432649	-	-	95%		-

SEQ ID No. 850	2268.2	Contig56	p	432873	434093	-	-	99%	-
SEQ ID No. 849	2266.2	Contig56	p	434150	435448	-	-	99%	-
SEQ ID No. 848	2264.2	Contig56	p	435357	436682	-	-	96%	-
SEQ ID No. 217	1242.2	Contig56	p	436686	437339	-	-	99%	-
SEQ ID No. 215	1239.2	Contig56	p	437445	437651	-	+		+
SEQ ID No. 214	1238.2	Contig56	p	437537	440545	-	-	97%	+
SEQ ID No. 2769	5188.2	Contig56	m	440810	442120	-	-	98%	-
SEQ ID No. 2770	5189.2	Contig56	m	442423	442962	+	-	100%	+
SEQ ID No. 1478	3235.3	Contig56	p	443439	444020	-	-	95%	-
SEQ ID No. 3109	5947.1	Contig56	p	443846	444142	-	+		-
SEQ ID No. 1479	3236.3	Contig56	m	444206	446461	-	-	98%	+
SEQ ID No. 3401	926.4	Contig56	m	446531	447229	-	-	97%	-
SEQ ID No. 3402	928.1	Contig56	m	447359	448597	-	-	97%	-
SEQ ID No. 3403	929.2	Contig56	m	448649	450079	-	-	97%	+
SEQ ID No. 573	1788.3	Contig56	m	450084	451466	-	-	98%	-
SEQ ID No. 575	1792.2	Contig56	m	451509	452315	-	-	99%	-
SEQ ID No. 1480	3238.1	Contig56	m	452294	453229	-	-	95%	-
SEQ ID No. 137	1116.2	Contig56	m	453233	453562	+	-	98%	+
SEQ ID No. 138	1118.2	Contig56	m	453914	455338	-	-	97%	-
SEQ ID No. 139	1119.2	Contig56	m	455395	456072	-	-	92%	+
SEQ ID No. 1481	3239.1	Contig56	p	456141	456452	+	-	100%	-
SEQ ID No. 1483	3240.3	Contig56	m	456543	456878	-	-	100%	-
SEQ ID No. 2118	4192.2	Contig56	p	456954	458858	-	-	97%	-
SEQ ID No. 2117	4191.1	Contig56	p	458774	460276	-	-	98%	-
SEQ ID No. 2115	4188.1	Contig56	p	460258	461391	-	-	95%	-
SEQ ID No. 2114	4186.1	Contig56	p	461427	462320	-	-	98%	-
SEQ ID No. 2113	4184.1	Contig56	p	462305	462637	-	-	99%	-
SEQ ID No. 477	1647.2	Contig56	m	462634	463245	-	-	98%	-
SEQ ID No. 476	1645.1	Contig56	p	463276	463866	-	-	98%	-
SEQ ID No. 475	1644.6	Contig56	m	463899	465158	-	-	95%	-
SEQ ID No. 1061	2583.3	Contig56	p	465296	465898	-	-	94%	-
SEQ ID No. 1060	2582.3	Contig56	p	465798	468422	-	-	97%	-
SEQ ID No. 2549	4809.2	Contig56	p	468423	469292	-	-	98%	-
SEQ ID No. 2551	4810.1	Contig56	p	469441	470094	-	-	98%	+
SEQ ID No. 2552	4811.2	Contig56	m	470268	470777	+	-	97%	-
SEQ ID No. 558	1768.3	Contig56	m	470901	472232	-	-	99%	-

SEQ ID No. 557	1767.2	Contig56	m	472211	472930	-	-	97%		-
SEQ ID No. 2562	4826.2	Contig56	m	472935	473642	+	-	99%		-
SEQ ID No. 2031	4073.1	Contig56	p	474278	475393	-	-	99%		-
SEQ ID No. 2030	4072.1	Contig56	p	475378	476064	-	-	98%		-
SEQ ID No. 3453	997.2	Contig56	p	476065	477177	-	-	99%		-
SEQ ID No. 3452	996.2	Contig56	p	477200	478231	-	-	98%		-
SEQ ID No. 3451	995.2	Contig56	p	478195	478866	-	-	91%		-
SEQ ID No. 2029	4071.1	Contig56	m	478863	479441	-	-	95%		+
SEQ ID No. 2028	4070.3	Contig56	p	479535	480974	-	-	99%		-
SEQ ID No. 2453	4678.3	Contig56	m	480971	481381		-	82%		+
SEQ ID No. 2455	4680.2	Contig56	m	481393	482142	-	-	98%		-
SEQ ID No. 2456	4683.3	Contig56	m	482069	483433	+	-	99%		-
SEQ ID No. 2914	5492.2	Contig56	m	484133	484753	+	-	99%		-
SEQ ID No. 2232	4356.2	Contig56	m	484717	485379	-	-	99%		-
SEQ ID No. 2231	4355.2	Contig56	m	485369	486358	-	-	98%	+	-
SEQ ID No. 2230	4354.2	Contig56	p	486613	486870	-	+			+
SEQ ID No. 2229	4353.2	Contig56	p	486710	487144		-	100%		+
SEQ ID No. 3303	787.3	Contig56	p	487168	489210	-	-	99%		-
SEQ ID No. 3304	789.4	Contig56	m	489432	489803	-	+			-
SEQ ID No. 2595	4880.2	Contig56	m	489804	490667	-	+			-
SEQ ID No. 2596	4881.4	Contig56	m	490946	494197	-	-	94%		-
SEQ ID No. 271	1331.3	Contig56	m	494404	495648	-	-	97%		-
SEQ ID No. 272	1332.3	Contig56	m	495853	497343	-	-	99%	+	-
SEQ ID No. 1749	3626.1	Contig56	m	497301	497801	-	-	99%	+	-
SEQ ID No. 1748	3625.2	Contig56	m	497609	498991	-	-	97%		-
SEQ ID No. 1747	3623.2	Contig56	m	498982	499392	-	-	100%		-
SEQ ID No. 1746	3622.3	Contig56	m	499393	500466	-	-	99%		-
SEQ ID No. 2653	4977.3	Contig56	p	500591	501079		-	100%		+
SEQ ID No. 2652	4974.1	Contig56	m	501054	501593	+	-	99%		-
SEQ ID No. 2536	4793.2	Contig56	p	502045	503820	-	-	99%		-
SEQ ID No. 2534	4790.3	Contig56	p	503795	505126	-	-	100%		-
SEQ ID No. 2535	4791.3	Contig56	m	504082	504483		-	100%		-
SEQ ID No. 2533	4788.3	Contig56	m	505287	505961	+	-	98%		-
SEQ ID No. 2656	4980.3	Contig56	m	505912	506535	+	-	99%		-
SEQ ID No. 2654	4979.1	Contig56	p	506583	507191	-	-	99%		-
SEQ ID No. 1265	2921.3	Contig56	p	507266	510616	-	-	96%		+

SEQ ID No. 1502	327.2	Contig56	m	510713	512743	-	-	97%		-
SEQ ID No. 1510	328.1	Contig56	m	512598	513521	-	-	98%		-
SEQ ID No. 1523	330.2	Contig56	m	513518	515047	-	-	99%	+	-
SEQ ID No. 1264	2920.1	Contig56	p	514374	514946		-	100%		+
SEQ ID No. 1262	2919.1	Contig56	m	515128	515355	-	+			-
SEQ ID No. 61	1003.3	Contig56	p	515479	517950	-	-	98%		+
SEQ ID No. 62	1005.2	Contig56	m	518054	518785	-	-	97%		-
SEQ ID No. 1261	2917.1	Contig56	p	519483	520199	-	-	97%		-
SEQ ID No. 1260	2916.1	Contig56	m	520040	520345		-	97%		+
SEQ ID No. 909	2360.2	Contig56	p	520200	521270	+	-	98%		-
SEQ ID No. 910	2361.4	Contig56	m	521370	527024	-	-	91%		+
SEQ ID No. 1707	3568.3	Contig56	m	527089	528615	-	-	99%		-
SEQ ID No. 1706	3566.1	Contig56	m	528497	529690	-	-	98%	+	-
SEQ ID No. 3251	712.2	Contig56	m	529698	530726	-	-	99%	+	-
SEQ ID No. 3250	711.2	Contig56	m	530771	532873	-	-	98%		-
SEQ ID No. 3249	709.1	Contig56	p	531595	531951		-	95%		+
SEQ ID No. 888	2335.2	Contig56	p	533008	534369	-	-	97%		+
SEQ ID No. 889	2336.2	Contig56	p	534412	536484	-	-	98%		-
SEQ ID No. 1705	3564.2	Contig56	m	536875	537150		-	100%		+
SEQ ID No. 1704	3563.2	Contig56	p	537264	537611		+			-
SEQ ID No. 3123	5988.1	Contig56	p	537425	537688		+			-

ORF Best-BlastP

10.1 Best-BlastP=> >nrprot 47% Identities = 35/90 (38%), Positives = 55/90 (61%), Gaps = 5/90 (5%) dbj|BAC94688.1| hypothetical protein [Vibrio vulnificus YJ016] Length = 343

1000.2 Best-BlastP=> >nrprot 62% Identities = 120/257 (46%), Positives = 162/257 (63%), Gaps = 3/257 (1%) ref|ZP_00079402.1| COG1024: Enoyl-CoA hydratase/carnithine racemase [Geobacter metallireducens] Length = 272

1001.3 Best-BlastP=> >nrprot 97% Identities = 419/437 (95%), Positives = 425/437 (97%) gb|AAM00614.1| chemiosmotic efflux system protein C-like protein [Legionella pneumophila] Length = 448

1002.2 Best-BlastP=> >nrprot 11% Identities = 35/129 (27%), Positives = 64/129 (49%), Gaps = 6/129 (4%) ref|ZP_00108772.1| COG0642: Signal transduction histidine kinase [Nostoc punctiforme] Length = 2053

1003.3 Best-BlastP=> >nrprot No Hits found

1005.2 Best-BlastP=> >nrprot 67% Identities = 126/243 (51%), Positives = 165/243 (67%), Gaps = 4/243 (1%) ref|NP_820690.1| dihydrodipicolinate reductase [Coxiella burnetii RSA 493] sp|P24703|DAPB_COXBU Dihydrodipicolinate reductase (DHPR) gb|AAO91204.1| dihydrodipicolinate reductase [Coxiella burnetii RSA 493] Length = 239

1006.3 Best-BlastP=> >nrprot 61% Identities = 96/211 (45%), Positives = 135/211 (63%), Gaps = 18/211 (8%) ref|NP_520174.1| PROBABLE TRANSMEMBRANE NADH DEHYDROGENASE I (CHAIN J) OXIDOREDUCTASE PROTEIN [Ralstonia solanacearum] emb|CAD15760.1| PROBABLE TRANSMEMBRANE NADH DEHYDROGENASE I (CHAIN J) OXIDOREDUCTASE PROTEIN [Ralstonia solanacearum] Length = 210

1009.2 Best-BlastP=> >nrprot 79% Identities = 113/159 (71%), Positives = 132/159 (83%) ref|NP_297604.1| NADH-ubiquinone oxidoreductase, NQO9 subunit [Xylella fastidiosa 9a5c] ref|ZP_00039600.1| hypothetical protein [Xylella fastidiosa Dixon] ref|ZP_00041894.1| hypothetical protein [Xylella fastidiosa Ann-1] ref|NP_778493.1| NADH-ubiquinone oxidoreductase NQO9 subunit [Xylella fastidiosa Temecula1] pir||C82822 NADH-ubiquinone oxidoreductase, NQO9 subunit XF0313 [imported] - Xylella fastidiosa (strain 9a5c) gb|AAF83124.1|AE003884_9 NADH-ubiquinone oxidoreductase, NQO9 subunit [Xylella fastidiosa 9a5c] gb|AAO28142.1| NADH-ubiquinone oxidoreductase NQO9 subunit [Xylella fastidiosa Temecula1] Length = 163

1010.2 Best-BlastP=> >nrprot 82% Identities = 223/332 (67%), Positives = 283/332 (85%) ref|NP_820424.1| NADH dehydrogenase I, H subunit [Coxiella burnetii RSA 493] gb|AAO90938.1| NADH dehydrogenase I, H subunit [Coxiella burnetii RSA 493] Length = 340

1011.5 Best-BlastP=> >nrprot 65% Identities = 374/800 (46%), Positives = 511/800 (63%), Gaps = 32/800 (4%) ref|NP_820425.1| NADH dehydrogenase I, G subunit [Coxiella burnetii RSA 493] gb|AAO90939.1| NADH dehydrogenase I, G subunit [Coxiella burnetii RSA 493] Length = 787

1013.1 Best-BlastP=> >nrprot No Hits found

1014.2 Best-BlastP=> >nrprot 55% Identities = 163/485 (33%), Positives = 254/485 (52%), Gaps = 36/485 (7%) ref|NP_711514.1| putative amidase [Leptospira interrogans serovar lai str. 56601] gb|AAN48532.1|AE011313_8 putative amidase [Leptospira interrogans serovar lai str. 56601] Length = 500

102.1 Best-BlastP=> >nrprot 92% Identities = 811/842 (96%), Positives = 830/842 (98%) gb|AAM00631.1| putative cation efflux transporter [Legionella pneumophila] Length = 842

1021.3 Best-BlastP=> >nrprot 76% Identities = 588/1008 (58%), Positives = 775/1008 (76%), Gaps = 4/1008 (0%) ref|NP_820092.1| transporter, AcrB/AcrD/AcrF family [Coxiella burnetii RSA 493] gb|AAO90606.1| transporter, AcrB/AcrD/AcrF family [Coxiella burnetii RSA 493] Length = 1019

1022.2 Best-BlastP=> >nrprot 47% Identities = 76/278 (27%), Positives = 135/278 (48%), Gaps = 6/278 (2%) ref|NP_655937.1| Acetyltransf, Acetyltransferase (GNAT) family [Bacillus anthracis A2012] ref|NP_844481.1| acetyltransferase, GNAT family [Bacillus anthracis str. Ames] gb|AAP25967.1| acetyltransferase, GNAT family [Bacillus anthracis str. Ames] Length = 295

1023.2 Best-BlastP=> >nrprot 65% Identities = 89/178 (50%), Positives = 126/178 (70%), Gaps = 3/178 (1%) ref|NP_799717.1| peptide methionine sulfoxide reductase [Vibrio parahaemolyticus RIMD 2210633] dbj|BAC61550.1| peptide methionine sulfoxide reductase [Vibrio parahaemolyticus] Length = 380

1024.2

Best-BlastP=> >nrprot 54% Identities = 71/193 (36%), Positives = 108/193 (55%), Gaps = 2/193 (1%) ref|NP_561497.1| conserved hypothetical protein [Clostridium perfringens] dbj|BAB80287.1| conserved hypothetical protein [Clostridium perfringens str. 13] Length = 195

1028.3 Best-BlastP=> >nrprot 67% Identities = 221/382 (57%), Positives = 274/382 (71%), Gaps = 2/382 (0%) ref|ZP_00032722.1| COG1228: Imidazolonepropionase and related amidohydrolases [Burkholderia fungorum] Length = 431

103.1 Best-BlastP=> >nrprot 99% Identities = 414/419 (98%), Positives = 418/419 (99%) gb|AAM00630.1| chemiosmotic efflux system B protein C [Legionella pneumophila] Length = 419

1030.3

Best-BlastP=> >nrprot 45% Identities = 70/234 (29%), Positives = 115/234 (49%), Gaps = 6/234 (2%) ref|NP_867781.1| probable sepiapterin reductase homolog yueD [Pirellula sp.] emb|CAD75328.1| probable sepiapterin reductase homolog yueD [Pirellula sp.] Length = 261

1032.3 Best-BlastP=> >nrprot 58% Identities = 131/317 (41%), Positives = 188/317 (59%), Gaps = 16/317 (5%) ref|NP_520766.1| PROBABLE ARGINASE PROTEIN [Ralstonia solanacearum] emb|CAD16352.1| PROBABLE ARGINASE PROTEIN [Ralstonia solanacearum] Length = 325

1034.3 Best-BlastP=> >nrprot 21% Identities = 43/131 (32%), Positives = 69/131 (52%), Gaps = 8/131 (6%) gb|AAL78307.1|AF288617_4 DotI [Legionella longbeachae] Length = 212

1035.3 Best-BlastP=> >nrprot 43% Identities = 124/492 (25%), Positives = 207/492 (42%), Gaps = 95/492 (19%) ref|NP_487180.1| hypothetical protein [Nostoc sp. PCC 7120] pir||AE2198 hypothetical protein alr3140 [imported] - Nostoc sp. (strain PCC 7120) dbj|BAB74839.1| ORF_ID:alr3140~hypothetical protein [Nostoc sp. PCC 7120] Length = 455

1036.4 Best-BlastP=> >nrprot 61% Identities = 206/421 (48%), Positives = 258/421 (61%), Gaps = 8/421 (1%) ref|ZP_00043700.1| COG0389: Nucleotidyltransferase/DNA polymerase involved in DNA repair [Magnetococcus sp. MC-1] Length = 421

1037.1 Best-BlastP=> >nrprot 68% Identities = 71/123 (57%), Positives = 96/123 (78%) ref|ZP_00043701.1| COG1974: SOS-response transcriptional repressors (RecA-mediated autopeptidases) [Magnetococcus sp. MC-1] Length = 238

1039.2 Best-BlastP=> >nrprot 59% Identities = 177/396 (44%), Positives = 241/396 (60%), Gaps = 5/396 (1%) ref|ZP_00094397.1| COG0624: Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases [Novosphingobium aromaticivorans] Length = 465

104.1 Best-BlastP=> >nrprot 99% Identities = 371/377 (98%), Positives = 376/377 (99%) gb|AAM00629.1| chemiosmotic efflux system B protein B [Legionella pneumophila] Length = 377

1040.3 Best-BlastP=> >nrprot 63% Identities = 150/366 (40%), Positives = 220/366 (60%), Gaps = 25/366 (6%) ref|NP_795120.1| conserved hypothetical protein [Pseudomonas syringae pv. tomato str. DC3000] gb|AAO58815.1| conserved hypothetical protein [Pseudomonas syringae pv. tomato str. DC3000] Length = 368

1041.3 Best-BlastP=> >nrprot 82% Identities = 557/801 (69%), Positives = 663/801 (82%), Gaps = 1/801 (0%) ref|NP_756480.1| DNA gyrase subunit B [Escherichia coli CFT073] gb|AAN83054.1|AE016769_169 DNA gyrase subunit B [Escherichia coli CFT073] Length = 805

1042.5 Best-BlastP=> >nrprot 76% Identities = 264/406 (65%), Positives = 314/406 (77%) ref|NP_719647.1| malate oxidoreductase, putative [Shewanella oneidensis MR-1] gb|AAN57091.1|AE015843_3 malate oxidoreductase, putative [Shewanella oneidensis MR-1] Length = 414

1044.3 Best-BlastP=> >nrprot 59% Identities = 152/394 (38%), Positives = 236/394 (59%), Gaps = 12/394 (3%) ref|NP_820159.1| major facilitator family transporter [Coxiella burnetii RSA 493] gb|AAO90673.1| major facilitator family transporter [Coxiella burnetii RSA 493] Length = 443

1047.2 Best-BlastP=> >nrprot 59% Identities = 142/261 (54%), Positives = 179/261 (68%), Gaps = 1/261 (0%) ref|ZP_00102270.1| COG1075: Predicted acetyltransferases and hydrolases with the alpha/beta hydrolase fold [Desulfitobacterium hafniense] Length = 364

1048.3 Best-BlastP=> >nrprot 12% Identities = 33/98 (33%), Positives = 48/98 (48%), Gaps = 3/98 (3%) ref|NP_717311.1| hypothetical protein [Shewanella oneidensis MR-1] gb|AAN54755.1|AE015616_1 hypothetical protein [Shewanella oneidensis MR-1] Length = 443

1049.2 Best-BlastP=> >nrprot 48% Identities = 74/232 (31%), Positives = 112/232 (48%), Gaps = 10/232 (4%) ref|NP_069069.1| glutamine ABC transporter, periplasmic glutamine-binding protein (glnH) [Archaeoglobus fulgidus DSM 4304] pir||G69278 glutamine ABC transporter, periplasmic glutamine-binding protein (glnH) homolog - Archaeoglobus fulgidus gb|AAB91001.1| glutamine ABC transporter, periplasmic glutamine-binding protein (glnH) [Archaeoglobus fulgidus DSM 4304] Length = 264

1050.4 Best-BlastP=> >nrprot 14% Identities = 96/410 (23%), Positives = 164/410 (40%), Gaps = 57/410 (13%) pir||T31688 Ca2+-transporting ATPase (EC 3.6.3.8), plasma membrane - Paramecium tetraurelia gb|AAB81284.1| plasma membrane calcium ATPase [Paramecium tetraurelia] Length = 1160

1053.2 Best-BlastP=> >nrprot 64% Identities = 86/174 (49%), Positives = 114/174 (65%), Gaps = 2/174 (1%) ref|NP_820971.1| phosphatase, putative [Coxiella burnetii RSA 493] gb|AAO91485.1| phosphatase, putative [Coxiella burnetii RSA 493] Length = 184

1055.1 Best-BlastP=> >nrprot 59% Identities = 102/256 (39%), Positives = 144/256 (56%), Gaps = 16/256 (6%) ref|NP_820853.1| conserved hypothetical protein [Coxiella burnetii RSA 493] gb|AAO91367.1| conserved hypothetical protein [Coxiella burnetii RSA 493] Length = 254

1056.1 Best-BlastP=> >nrprot No Hits found

1058.3 Best-BlastP=> >nrprot 43% Identities = 112/511 (21%), Positives = 204/511 (39%), Gaps = 94/511 (18%) sp|P08799|MYS2_DICDI Myosin II heavy chain, non muscle pir||A26655 myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum) gb|AAA33227.1| myosin heavy chain Length = 2116

1059.2 Best-BlastP=> >nrprot 85% Identities = 302/423 (71%), Positives = 364/423 (86%), Gaps = 2/423 (0%) ref|NP_820394.1| citrate synthase [Coxiella burnetii RSA 493] sp|P18789|CISY_COXBU Citrate synthase pir||JQ1392 citrate (si)-synthase (EC 4.1.3.7) - Coxiella burnetii gb|AAA23307.1| citrate synthase (g1tA) (EC 4.1.3.7) gb|AAO90908.1| citrate synthase [Coxiella burnetii RSA 493] Length = 430

1060.2 Best-BlastP=> >nrprot 57% Identities = 114/286 (39%), Positives = 168/286 (58%), Gaps = 7/286 (2%) ref|NP_813465.1| purine nucleoside phosphorylase II [Bacteroides thetaiotaomicron VPI-5482] gb|AAO79659.1| purine nucleoside phosphorylase II [Bacteroides thetaiotaomicron VPI-5482] Length = 292

1063.1 Best-BlastP=> >nrprot No Hits found

1065.3 Best-BlastP=> >nrprot 48% Identities = 124/365 (33%), Positives = 187/365 (51%), Gaps = 10/365 (2%) dbj|BAA31547.1| metal-activated pyridoxal enzyme [Arthrobacter sp.] Length = 379

1066.2 Best-BlastP=> >nrprot 53% Identities = 124/294 (42%), Positives = 160/294 (54%), Gaps = 18/294 (6%) ref|ZP_00016110.1| COG0596: Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily) [Rhodospirillum rubrum] Length = 286

1067.1 Best-BlastP=> >nrprot 47% Identities = 28/70 (40%), Positives = 40/70 (57%), Gaps = 11/70 (15%) ref|NP_489272.1| unknown protein [Nostoc sp. PCC 7120] pir||AH2459 hypothetical protein alr5232 [imported] - Nostoc sp. (strain PCC 7120) dbj|BAB76931.1| ORF_ID:alr5232~unknown protein [Nostoc sp. PCC 7120] Length = 204

1069.2 Best-BlastP=> >nrprot 46% Identities = 103/273 (37%), Positives = 145/273 (53%), Gaps = 35/273 (12%) emb|CAA75849.1| hypothetical protein [Coxiella burnetii] Length = 309

107.1 Best-BlastP=> >nrprot 99% Identities = 1034/1047 (98%), Positives = 1041/1047 (99%) gb|AAM00628.1| chemiosmotic efflux system B protein A [Legionella pneumophila] Length = 1047

1071.3 Best-BlastP=> >nrprot 99% Identities = 411/418 (98%), Positives = 416/418 (99%) pir||T08882 proline/betaine transport protein homolog CitA - Legionella pneumophila emb|CAA75171.1| TphA protein [Legionella pneumophila] gb|AAC38182.1| CitA [Legionella pneumophila] emb|CAA75337.1| TphA protein [Legionella pneumophila] Length = 418

1072.4 Best-BlastP=> >nrprot 99% Identities = 966/973 (99%), Positives = 971/973 (99%) pir||T18341 icmF protein - Legionella pneumophila emb|CAA75172.1| IcmF protein [Legionella pneumophila] emb|CAA75338.1| IcmF protein [Legionella pneumophila] Length = 973

1074.3 Best-BlastP=> >nrprot 67% Identities = 343/638 (53%), Positives = 456/638 (71%), Gaps = 6/638 (0%) ref|NP_819606.1| fatty oxidation complex, alpha subunit [Coxiella burnetii RSA 493] gb|AAO90120.1| fatty oxidation complex, alpha subunit [Coxiella burnetii RSA 493] Length = 642

1075.3 Best-BlastP=> >nrprot 75% Identities = 271/422 (64%), Positives = 333/422 (78%), Gaps = 1/422 (0%) emb|CAD58320.1| beta-Subunit of fatty acid oxidation complex, 3-keto-acyl-CoA-thiolase [Azoarcus sp. EbN1] Length = 426

1076.3 Best-BlastP=> >nrprot 56% Identities = 50/119 (42%), Positives = 70/119 (58%) ref|NP_902011.1| conserved hypothetical protein [Chromobacterium violaceum ATCC 12472] gb|AAQ60013.1| conserved hypothetical protein [Chromobacterium violaceum ATCC 12472] Length = 122

1077.1 Best-BlastP=> >nrprot 75% Identities = 137/226 (60%), Positives = 175/226 (77%), Gaps = 1/226 (0%) ref|NP_384333.1| PUTATIVE TRANSCRIPTION REGULATOR PROTEIN [Sinorhizobium meliloti] emb|CAC41614.1| PUTATIVE TRANSCRIPTION REGULATOR PROTEIN [Sinorhizobium meliloti] Length = 230

1078.2 Best-BlastP=> >nrprot 33% Identities = 65/262 (24%), Positives = 110/262 (41%), Gaps = 21/262 (8%) ref|NP_624872.1| hypothetical protein SCF73.06c [Streptomyces coelicolor A3(2)] emb|CAB57411.1| hypothetical protein SCF73.06c [Streptomyces coelicolor A3(2)] Length = 333

1080.2 Best-BlastP=> >nrprot 40% Identities = 71/229 (31%), Positives = 113/229 (49%), Gaps = 11/229 (4%) ref|ZP_00081514.1| COG1560: Lauroyl/myristoyl acyltransferase [Geobacter metallireducens] Length = 324

1081.2 Best-BlastP=> >nrprot 36% Identities = 81/362 (22%), Positives = 140/362 (38%), Gaps = 52/362 (14%) ref|ZP_00133556.1| COG5295: Autotransporter adhesin [Haemophilus somnus 2336] Length = 3391

1082.3 Best-BlastP=> >nrprot 67% Identities = 236/481 (49%), Positives = 325/481 (67%), Gaps = 12/481 (2%) ref|NP_484210.1| alpha,alpha-trehalase [Nostoc sp. PCC 7120] pir||AF1827 alpha,alpha-trehalase [imported] - Nostoc sp. (strain PCC 7120) dbj|BAB77690.1| alpha,alpha-trehalase [Nostoc sp. PCC 7120] Length = 495

1087.2 Best-BlastP=> >nrprot 56% Identities = 191/581 (32%), Positives = 316/581 (54%), Gaps = 34/581 (5%) ref|NP_488098.1| similar to isovaleryl-CoA dehydrogenase [Nostoc sp. PCC 7120] pir||AC2313 hypothetical protein alr4058 [imported] - Nostoc sp. (strain PCC 7120) dbj|BAB75757.1| ORF_ID:alr4058~similar to isovaleryl-CoA dehydrogenase [Nostoc sp. PCC 7120] Length = 602

1088.3 Best-BlastP=> >nrprot 40% Identities = 124/450 (27%), Positives = 202/450 (44%), Gaps = 37/450 (8%) ref|NP_870622.1| probable auxin-responsive-like protein [Pirellula sp.] emb|CAD77699.1| probable auxin-responsive-like protein [Pirellula sp.] Length = 559

1089.2 Best-BlastP=> >nrprot 48% Identities = 274/902 (30%), Positives = 444/902 (49%), Gaps = 54/902 (5%) ref|ZP_00123122.1| hypothetical protein [Haemophilus somnus 129PT] Length = 948

109.1 Best-BlastP=> >nrprot 96% Identities = 441/469 (94%), Positives = 456/469 (97%) gb|AAM00627.1| unknown [Legionella pneumophila] Length = 470

1090.1 Best-BlastP=> >nrprot 43% Identities = 28/84 (33%), Positives = 51/84 (60%), Gaps = 12/84 (14%) ref|ZP_00025260.1| hypothetical protein [Ralstonia metallidurans] Length = 241

1091.2 Best-BlastP=> >nrprot 45% Identities = 114/438 (26%), Positives = 192/438 (43%), Gaps = 79/438 (18%) ref|NP_873428.1| conserved hypothetical protein [Haemophilus ducreyi 35000HP] gb|AAP95817.1| conserved hypothetical protein [Haemophilus ducreyi 35000HP] Length = 481

1092.3 Best-BlastP=> >nrprot 46% Identities = 64/221 (28%), Positives = 115/221 (52%), Gaps = 8/221 (3%) ref|ZP_00087883.1| COG1940: Transcriptional regulator/sugar kinase [Pseudomonas fluorescens PfO-1] Length = 271

1093.3 Best-BlastP=> >nrprot 74% Identities = 161/325 (49%), Positives = 232/325 (71%), Gaps = 16/325 (4%) ref|NP_819394.1| sohB protein [Coxiella burnetii RSA 493] gb|AAO89908.1| sohB protein [Coxiella burnetii RSA 493] Length = 338

1094.1 Best-BlastP=> >nrprot 36% Identities = 30/50 (60%), Positives = 34/50 (68%), Gaps = 4/50 (8%) gb|EAA17424.1| putative HSP protein [Plasmodium yoelii yoelii] Length = 929

1096.3 Best-BlastP=> >nrprot 70% Identities = 253/483 (52%), Positives = 333/483 (68%), Gaps = 17/483 (3%) ref|ZP_00127277.1| COG2204: Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains [Pseudomonas syringae pv. syringae B728a] Length = 491

1098.3 Best-BlastP=> >nrprot 57% Identities = 136/330 (41%), Positives = 195/330 (59%), Gaps = 3/330 (0%) ref|NP_718989.1| conserved hypothetical protein [Shewanella oneidensis MR-1] sp|Q8EBR4|TRUD_SHEON tRNA pseudouridine synthase D (Pseudouridylate synthase) (Uracil hydrolyase) gb|AAN56433.1|AE015780_4 conserved hypothetical protein [Shewanella oneidensis MR-1] Length = 370

1099.2 Best-BlastP=> >nrprot 63% Identities = 158/328 (48%), Positives = 207/328 (63%), Gaps = 6/328 (1%) ref|NP_403834.1| conserved hypothetical protein [Yersinia pestis] ref|NP_671254.1| hypothetical protein [Yersinia pestis KIM] pir||AH0022 conserved hypothetical protein YPO0180 [imported] - Yersinia pestis (strain CO92) emb|CAC89042.1| conserved hypothetical protein [Yersinia pestis CO92] gb|AAM87505.1|AE014000_9 hypothetical protein [Yersinia pestis KIM] Length = 326

11.2 Best-BlastP=> >nrprot 39% Identities = 77/276 (27%), Positives = 127/276 (46%), Gaps = 29/276 (10%) ref|ZP_00023112.1| hypothetical protein [Ralstonia metallidurans] Length = 348

1101.4 Best-BlastP=> >nrprot 54% Identities = 128/359 (35%), Positives = 200/359 (55%), Gaps = 24/359 (6%) ref|NP_878747.1| histidinol-phosphate aminotransferase [Candidatus Blochmannia floridanus] emb|CAD83523.1| histidinol-phosphate aminotransferase [Candidatus Blochmannia floridanus] Length = 356

1102.2 Best-BlastP=> >nrprot 66% Identities = 203/427 (47%), Positives = 287/427 (67%), Gaps = 9/427 (2%) dbj|BAC94115.1| histidinol dehydrogenase [Vibrio vulnificus YJ016] Length = 431

1105.2 Best-BlastP=> >nrprot 36% Identities = 187/439 (42%), Positives = 279/439 (63%), Gaps = 1/439 (0%) ref|NP_715981.1| sensory box protein [Shewanella oneidensis MR-1] gb|AAN53426.1|AE015481_9 sensory box protein [Shewanella oneidensis MR-1] Length = 1515

1106.5 Best-BlastP=> >nrprot 99% Identities = 1366/1368 (99%), Positives = 1367/1368 (99%) gb|AAC69338.1| RNA polymerase B-subunit [Legionella pneumophila] Length = 1368

1107.2 Best-BlastP=> >nrprot 66% Identities = 154/324 (47%), Positives = 219/324 (67%), Gaps = 3/324 (0%) ref|NP_657119.1| Chal_stil_syntC, Chalcone and stilbene synthases, C-terminal domain [Bacillus anthracis A2012] ref|NP_845551.1| 3-oxoacyl-(acyl-carrier-protein) synthase III, putative [Bacillus anthracis str. Ames] gb|AAP27037.1| 3-oxoacyl-(acyl-carrier-protein) synthase III, putative [Bacillus anthracis str. Ames] Length = 329

1109.3 Best-BlastP=> >nrprot 64% Identities = 150/329 (45%), Positives = 212/329 (64%), Gaps = 3/329 (0%) ref|NP_657117.1| 3Beta_HSD, 3-beta hydroxysteroid dehydrogenase/isomerase family [Bacillus anthracis A2012] ref|NP_845549.1| 3-beta hydroxysteroid dehydrogenase/isomerase family protein [Bacillus anthracis str. Ames] gb|AAP27035.1| 3-beta hydroxysteroid dehydrogenase/isomerase family protein [Bacillus anthracis str. Ames] Length = 328

111.2 Best-BlastP=> >nrprot 97% Identities = 485/504 (96%), Positives = 493/504 (97%) sp|Q8RNP4|TYPH_LEGPN Putative thymidine phosphorylase (TdRPase) gb|AAM00626.1| unknown [Legionella pneumophila] Length = 517

1111.3 Best-BlastP=> >nrprot 58% Identities = 109/270 (40%), Positives = 164/270 (60%), Gaps = 6/270 (2%) ref|NP_657116.1| hypothetical protein predicted by GeneMark [Bacillus anthracis A2012] ref|NP_845548.1| conserved hypothetical protein [Bacillus anthracis str. Ames] gb|AAP27034.1| conserved hypothetical protein [Bacillus anthracis str. Ames] Length = 283

1113.2 Best-BlastP=> >nrprot 31% Identities = 37/58 (63%), Positives = 42/58 (72%), Gaps = 3/58 (5%) ref|NP_759576.1| Conserved hypothetical protein [Vibrio vulnificus CMCP6] gb|AAO09103.1|AE016799_1 Conserved hypothetical protein [Vibrio vulnificus CMCP6] Length = 116

1115.4 Best-BlastP=> >nrprot 26% Identities = 41/132 (31%), Positives = 70/132 (53%), Gaps = 10/132 (7%) ref|NP_710625.1| conserved hypothetical protein [Leptospira interrogans serovar lai str. 56601] gb|AAN47643.1|AE011231_6 conserved hypothetical protein [Leptospira interrogans serovar lai str. 56601] Length = 359

1116.2 Best-BlastP=> >nrprot No Hits found

1118.2 Best-BlastP=> >nrprot 79% Identities = 277/452 (61%), Positives = 359/452 (79%), Gaps = 1/452 (0%) gb|AAM00627.1| unknown [Legionella pneumophila] Length = 470

1119.2 Best-BlastP=> >nrprot No Hits found

1122.1 Best-BlastP=> >nrprot 64% Identities = 142/291 (48%), Positives = 196/291 (67%), Gaps = 5/291 (1%) ref|NP_716239.1| hflC protein [Shewanella oneidensis MR-1] gb|AAN53684.1|AE015507_10 hflC protein [Shewanella oneidensis MR-1] Length = 297

1123.2 Best-BlastP=> >nrprot 62% Identities = 176/379 (46%), Positives = 237/379 (62%), Gaps = 24/379 (6%) ref|NP_253629.1| protease subunit HflK [Pseudomonas aeruginosa PA01] pir||B83028 proteinase subunit HflK PA4942 [imported] - Pseudomonas aeruginosa (strain PAO1) gb|AAG08327.1|AE004907_5 protease subunit HflK [Pseudomonas aeruginosa PAO1] Length = 400

1124.2 Best-BlastP=> >nrprot 36% Identities = 58/152 (38%), Positives = 87/152 (57%), Gaps = 8/152 (5%) gb|AAM09314.2| similar to Mus musculus (Mouse). Uridine-cytidine kinase 2 [Dictyostelium discoideum] Length = 243

1125.2 Best-BlastP=> >nrprot 68% Identities = 193/382 (50%), Positives = 265/382 (69%) ref|NP_519762.1| PROBABLE ACETYLORNITHINE DEACETYLASE (ACETYLORNITHINASE) PROTEIN [Ralstonia solanacearum] emb|CAD15343.1| PROBABLE ACETYLORNITHINE DEACETYLASE (ACETYLORNITHINASE) PROTEIN [Ralstonia solanacearum] Length = 397

1126.3 Best-BlastP=> >nrprot 55% Identities = 163/398 (40%), Positives = 239/398 (60%), Gaps = 2/398 (0%) ref|ZP_00033611.1| COG0477: Permeases of the major facilitator superfamily [Burkholderia fungorum] Length = 470

1127.3 Best-BlastP=> >nrprot 72% Identities = 360/650 (55%), Positives = 472/650 (72%), Gaps = 4/650 (0%) ref|NP_745215.1| acetoacetyl-CoA synthetase, putative [Pseudomonas putida KT2440] gb|AAN68679.1|AE016497_4 acetoacetyl-CoA synthetase, putative [Pseudomonas putida KT2440] Length = 650

1128.3 Best-BlastP=> >nrprot 50% Identities = 113/289 (39%), Positives = 170/289 (58%), Gaps = 13/289 (4%) ref|NP_904046.1| biotin synthesis protein [Chromobacterium violaceum ATCC 12472] gb|AAQ62035.1| biotin synthesis protein [Chromobacterium violaceum ATCC 12472] Length = 302

1132.2 Best-BlastP=> >nrprot 57% Identities = 85/221 (38%), Positives = 135/221 (61%), Gaps = 12/221 (5%) ref|ZP_00086846.1| COG1040: Predicted amidophosphoribosyltransferases [Pseudomonas fluorescens PfO-1] Length = 246

1133.3 Best-BlastP=> >nrprot 50% Identities = 103/346 (29%), Positives = 172/346 (49%), Gaps = 13/346 (3%) ref|ZP_00082200.1| COG0750: Predicted membrane-associated Zn-dependent proteases 1 [Geobacter metallireducens] Length = 355

1134.3 Best-BlastP=> >nrprot No Hits found

1136.3 Best-BlastP=> >nrprot 62% Identities = 232/442 (52%), Positives = 307/442 (69%), Gaps = 1/442 (0%) ref|ZP_00043253.1| COG0790: FOG: TPR repeat, SEL1 subfamily [Magnetococcus sp. MC-1] Length = 831

1137.2 Best-BlastP=> >nrprot No Hits found

114.2 Best-BlastP=> >nrprot No Hits found

1142.4 Best-BlastP=> >nrprot 65% Identities = 222/454 (48%), Positives = 312/454 (68%), Gaps = 4/454 (0%) ref|NP_820336.1| amino acid antiporter [Coxiella burnetii RSA 493] gb|AAO90850.1| amino acid antiporter [Coxiella burnetii RSA 493] Length = 474

1144.3 Best-BlastP=> >nrprot 64% Identities = 180/379 (47%), Positives = 254/379 (67%), Gaps = 4/379 (1%) ref|NP_800676.1| conserved hypothetical protein [Vibrio parahaemolyticus RIMD 2210633] sp|Q87GZ9|CLCA_VIBPA Voltage-gated ClC-type chloride channel clcA dbj|BAC62509.1| conserved hypothetical protein [Vibrio parahaemolyticus] Length = 467

1145.3 Best-BlastP=> >nrprot 72% Identities = 391/721 (54%), Positives = 522/721 (72%), Gaps = 11/721 (1%) ref|ZP_00110122.1| COG0512: Anthranilate/para-aminobenzoate synthases component II [Nostoc punctiforme] Length = 734

1148.2

Best-BlastP=> >nrprot 56% Identities = 334/906 (36%), Positives = 512/906 (56%), Gaps = 35/906 (3%) ref|NP_289628.1| adenylylating enzyme for glutamine synthetase [Escherichia coli O157:H7 EDL933] ref|NP_311963.1| glutamate-ammonia-ligase adenylyltransferase [Escherichia coli O157:H7] pir||H91120 glutamate-ammonia-ligase adenylyltransferase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952) pir||G85965 adenylylating enzyme for glutamine synthetase [imported] - Escherichia coli (strain O157:H7, substrain EDL933) gb|AAG58187.1|AE005534_9 adenylylating enzyme for glutamine synthetase [Escherichia coli O157:H7 EDL933] dbj|BAB37359.1| glutamate-ammonia-ligase adenylyltransferase [Escherichia coli O157:H7] Length = 946

115.2 Best-BlastP=> >nrprot 57% Identities = 53/167 (31%), Positives = 98/167 (58%), Gaps = 5/167 (2%) ref|NP_490158.1| probable acetyltransferase [Nostoc sp. PCC 7120] pir||AD2484 hypothetical protein alr7052 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha dbj|BAB78136.1| ORF_ID:alr7052~probable acetyltransferase [Nostoc sp. PCC 7120] Length = 169

1152.2 Best-BlastP=> >nrprot 22% Identities = 39/137 (28%), Positives = 67/137 (48%), Gaps = 10/137 (7%) ref|NP_824638.1| putative lipase [Streptomyces avermitilis MA-4680] dbj|BAC71173.1| putative lipase [Streptomyces avermitilis MA-4680] Length = 286

1153.1 Best-BlastP=> >nrprot 46% Identities = 36/136 (26%), Positives = 67/136 (49%), Gaps = 15/136 (11%) ref|NP_629974.1| putative integral membrane protein [Streptomyces coelicolor A3(2)] pir||T35887 hypothetical protein SC9B10.18 - Streptomyces coelicolor emb|CAA15808.1| putative integral membrane protein [Streptomyces coelicolor A3(2)] Length = 312

1156.3 Best-BlastP=> >nrprot 48% Identities = 114/437 (26%), Positives = 213/437 (48%), Gaps = 37/437 (8%) ref|ZP_00107812.1| COG4325: Predicted membrane protein [Nostoc punctiforme] Length = 449

1157.4 Best-BlastP=> >nrprot 73% Identities = 550/943 (58%), Positives = 687/943 (72%), Gaps = 15/943 (1%) ref|ZP_00085403.1| COG0567: 2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, and related enzymes [Pseudomonas fluorescens PfO-1] Length = 943

1159.5 Best-BlastP=> >nrprot 30% Identities = 61/269 (22%), Positives = 130/269 (48%), Gaps = 23/269 (8%) ref|NP_143635.1| chromosome assembly protein [Pyrococcus horikoshii] pir||F71190 probable chromosome assembly protein - Pyrococcus horikoshii dbj|BAA30917.1| 1179aa long hypothetical chromosome assembly protein [Pyrococcus horikoshii] Length = 1179

116.1 Best-BlastP=> >nrprot 29% Identities = 42/123 (34%), Positives = 47/123 (38%), Gaps = 27/123 (21%) ref|NP_639328.1| hypothetical protein [Xanthomonas campestris pv. campestris str. ATCC 33913] gb|AAM43210.1| hypothetical protein [Xanthomonas campestris pv. campestris str. ATCC 33913] Length = 131

1160.2 Best-BlastP=> >nrprot 51% Identities = 136/392 (34%), Positives = 206/392 (52%), Gaps = 10/392 (2%) ref|NP_085189.1| IS10 orf [Shigella flexneri] ref|NP_858160.1| hypothetical protein [Shigella flexneri 2a] gb|AAK18345.1|AF348706_34 IS10 orf [Shigella flexneri] gb|AAL72480.1| hypothetical protein [Shigella flexneri 2a] Length = 407

1161.2 Best-BlastP=> >nrprot 75% Identities = 225/353 (63%), Positives = 269/353 (76%), Gaps = 2/353 (0%) ref|NP_756808.1| Uroporphyrinogen decarboxylase [Escherichia coli CFT073] sp|Q8FB74|DCUP_ECOL6 Uroporphyrinogen decarboxylase (URO-D) (UPD) gb|AAN83382.1|AE016770_182 Uroporphyrinogen decarboxylase [Escherichia coli CFT073] Length = 354

1162.1 Best-BlastP=> >nrprot 62% Identities = 47/114 (41%), Positives = 71/114 (62%), Gaps = 2/114 (1%) ref|NP_289633.1| putative kinase [Escherichia coli O157:H7 EDL933] ref|NP_311968.1| putative kinase [Escherichia coli O157:H7] pir||E91121 probable kinase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952) pir||D85966 probable kinase ygiG [imported] - Escherichia coli (strain O157:H7, substrain EDL933) gb|AAG58192.1|AE005535_4 putative kinase [Escherichia coli O157:H7 EDL933] dbj|BAB37364.1| putative kinase [Escherichia coli O157:H7] Length = 123

1163.2 Best-BlastP=> >nrprot 29% Identities = 46/156 (29%), Positives = 68/156 (43%), Gaps = 10/156 (6%) ref|NP_747191.1| conserved hypothetical protein [Pseudomonas putida KT2440] gb|AAN70655.1|AE016709_7 conserved hypothetical protein [Pseudomonas putida KT2440] Length = 232

1168.2 Best-BlastP=> >nrprot 61% Identities = 64/117 (54%), Positives = 83/117 (70%) ref|NP_755691.1| Hypothetical protein [Escherichia coli CFT073] gb|AAN82265.1|AE016767_25 Hypothetical protein [Escherichia coli CFT073] Length = 365

1169.3 Best-BlastP=> >nrprot 99% Identities = 616/621 (99%), Positives = 618/621 (99%), Gaps = 1/621 (0%) gb|AAB09543.1| LprpoD Length = 620

117.1 Best-BlastP=> >nrprot No Hits found

1171.3 Best-BlastP=> >nrprot 64% Identities = 104/241 (43%), Positives = 161/241 (66%), Gaps = 6/241 (2%) sp|P13980|TRT3_ECOLI TraT complement resistance protein precursor pir||C29835 TraTp protein - Escherichia coli plasmid pED208 gb|AAA88375.1| traTp gene product gb|AAM90725.2| TraT [Salmonella typhi] Length = 245

1172.2 Best-BlastP=> >nrprot No Hits found

1174.2 Best-BlastP=> >nrprot 74% Identities = 267/451 (59%), Positives = 334/451 (74%), Gaps = 9/451 (1%) ref|ZP_00065505.1| COG0486: Predicted GTPase [Microbulbifer degradans 2-40] Length = 456

1176.5 Best-BlastP=> >nrprot No Hits found

1179.1 Best-BlastP=> >nrprot No Hits found

118.1 Best-BlastP=> >nrprot No Hits found

1180.2 Best-BlastP=> >nrprot 44% Identities = 66/232 (28%), Positives = 113/232 (48%), Gaps = 2/232 (0%) ref|NP_762946.1| AraC-type DNA-binding domain-containing protein [Vibrio vulnificus CMCP6] gb|AAO07936.1|AE016811_177 AraC-type DNA-binding domain-containing protein [Vibrio vulnificus CMCP6] Length = 237

1182.3 Best-BlastP=> >nrprot 48% Identities = 88/274 (32%), Positives = 149/274 (54%), Gaps = 18/274 (6%) ref|NP_520074.1| PROBABLE TRANSMEMBRANE PROTEIN [Ralstonia solanacearum] emb|CAD15655.1| PROBABLE TRANSMEMBRANE PROTEIN [Ralstonia solanacearum] Length = 336

1183.4 Best-BlastP=> >nrprot No Hits found

1184.2

Best-BlastP=> >nrprot 54% Identities = 133/395 (33%), Positives = 216/395 (54%), Gaps = 6/395 (1%) ref|NP_618268.1| amino acid transporter [Methanosarcina acetivorans str. C2A] gb|AAM06748.1| amino acid transporter [Methanosarcina acetivorans str. C2A] Length = 439

1186.2 Best-BlastP=> >nrprot 31% Identities = 52/171 (30%), Positives = 82/171 (47%), Gaps = 7/171 (4%) ref|NP_149698.1| 235L [Invertebrate iridescent virus 6] gb|AAK82096.1|AF303741_235 235L [Chilo iridescent virus] Length = 265

1188.2 Best-BlastP=> >nrprot 99% Identities = 737/749 (98%), Positives = 746/749 (99%) sp|Q9WXB9|CATA_LEGPN Peroxidase/catalase (Catalase-peroxidase) dbj|BAA78342.1| catalase-peroxidase [Legionella pneumophila] gb|AAG37106.1|AF276752_1 catalase-peroxidase [Legionella pneumophila] Length = 749

119.1 Best-BlastP=> >nrprot 51% Identities = 53/124 (42%), Positives = 77/124 (62%), Gaps = 2/124 (1%) dbj|BAA75251.1| Similar to IS1301 of Neisseria meningitidis [Actinobacillus actinomycetemcomitans] Length = 255

1190.4 Best-BlastP=> >nrprot 56% Identities = 256/709 (36%), Positives = 406/709 (57%), Gaps = 25/709 (3%) ref|ZP_00065821.1| COG1249: Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes [Microbulbifer degradans 2-40] Length = 704

1193.2 Best-BlastP=> >nrprot No Hits found

1194.3 Best-BlastP=> >nrprot 72% Identities = 223/383 (58%), Positives = 278/383 (72%), Gaps = 1/383 (0%) ref|NP_636802.1| glutaryl-CoA dehydrogenase [Xanthomonas campestris pv. campestris str. ATCC 33913] gb|AAM40726.1| glutaryl-CoA dehydrogenase [Xanthomonas campestris pv. campestris str. ATCC 33913] Length = 387

1199.4 Best-BlastP=> >nrprot 80% Identities = 73/110 (66%), Positives = 90/110 (81%) ref|ZP_00067987.1| COG0091: Ribosomal protein L22 [Microbulbifer degradans 2-40] Length = 110

1201.1 Best-BlastP=> >nrprot 82% Identities = 150/220 (68%), Positives = 180/220 (81%), Gaps = 1/220 (0%) ref|NP_819288.1| ribosomal protein S3 [Coxiella burnetii RSA 493] sp|O85388|RS3_COXBU 30S ribosomal protein S3 gb|AAO89802.1| ribosomal protein S3 [Coxiella burnetii RSA 493] Length = 227

1202.2 Best-BlastP=> >nrprot 89% Identities = 113/137 (82%), Positives = 124/137 (90%) ref|NP_742627.1| ribosomal protein L16 [Pseudomonas putida KT2440] gb|AAN66091.1|AE016238_9 ribosomal protein L16 [Pseudomonas putida KT2440] Length = 137

1205.5 Best-BlastP=> >nrprot 16% Identities = 428/1182 (36%), Positives = 630/1182 (53%), Gaps = 81/1182 (6%) ref|ZP_00108802.1| hypothetical protein [Nostoc punctiforme] Length = 2315

1207.2 Best-BlastP=> >nrprot No Hits found

1208.1 Best-BlastP=> >nrprot 42% Identities = 45/124 (36%), Positives = 55/124 (44%), Gaps = 4/124 (3%) pir||F71456 hypothetical protein PH0308 - Pyrococcus horikoshii dbj|BAA29381.1| 215aa long hypothetical protein [Pyrococcus horikoshii] Length = 215

1209.2 Best-BlastP=> >nrprot 84% Identities = 346/487 (71%), Positives = 417/487 (85%), Gaps = 1/487 (0%) ref|ZP_00087773.1| COG0516: IMP dehydrogenase/GMP reductase [Pseudomonas fluorescens PfO-1] Length = 506

121.1 Best-BlastP=> >nrprot 59% Identities = 58/101 (57%), Positives = 67/101 (66%), Gaps = 1/101 (0%) ref|ZP_00091135.1| COG2852: Uncharacterized protein conserved in bacteria [Azotobacter vinelandii] Length = 150

1210.2 Best-BlastP=> >nrprot 81% Identities = 354/523 (67%), Positives = 427/523 (81%), Gaps = 1/523 (0%) ref|NP_708346.1| GMP synthetase [Shigella flexneri 2a str. 301] ref|NP_838070.1| GMP synthetase (glutamine-hydrolyzing) [Shigella flexneri 2a str. 2457T] gb|AAN44053.1|AE015270_8 GMP synthetase [Shigella flexneri 2a str. 301] gb|AAP17880.1| GMP synthetase (glutamine-hydrolyzing) [Shigella flexneri 2a str. 2457T] Length = 525

1211.3 Best-BlastP=> >nrprot 19% Identities = 47/90 (52%), Positives = 69/90 (76%), Gaps = 1/90 (1%) ref|NP_929783.1| hypothetical protein [Photorhabdus luminescens subsp. laumondii TTO1] emb|CAE14921.1| unnamed protein product [Photorhabdus luminescens subsp. laumondii TTO1] Length = 94

1213.2 Best-BlastP=> >nrprot 65% Identities = 193/394 (48%), Positives = 248/394 (62%), Gaps = 22/394 (5%) ref|NP_249543.1| chitin-binding protein CbpD precursor [Pseudomonas aeruginosa PA01] pir||F83538 chitin-binding protein CbpD precursor PA0852 [imported] - Pseudomonas aeruginosa (strain PAO1) gb|AAG04241.1|AE004520_4 chitin-binding protein CbpD precursor [Pseudomonas aeruginosa PAO1] Length = 389

1215.3 Best-BlastP=> >nrprot 64% Identities = 172/363 (47%), Positives = 233/363 (64%), Gaps = 5/363 (1%) ref|NP_387138.1| CONSERVED HYPOTHETICAL PROTEIN [Sinorhizobium meliloti] emb|CAC47611.1| CONSERVED HYPOTHETICAL PROTEIN [Sinorhizobium meliloti] Length = 367

122.3 Best-BlastP=> >nrprot 69% Identities = 90/159 (56%), Positives = 115/159 (72%), Gaps = 4/159 (2%) ref|NP_779515.1| chromosome partitioning protein [Xylella fastidiosa Temecula1] sp|Q87BY1|PARB_XYLFT Probable chromosome partitioning protein parB gb|AAO29164.1| chromosome partitioning protein [Xylella fastidiosa Temecula1] Length = 310

1220.3 Best-BlastP=> >nrprot 77% Identities = 332/507 (65%), Positives = 395/507 (77%), Gaps = 3/507 (0%) dbj|BAB19801.1| piperideine-6-carboxylate dehydrogenase ['Flavobacterium' lutescens] Length = 510

1222.5 Best-BlastP=> >nrprot No Hits found

1223.2 Best-BlastP=> >nrprot 39% Identities = 73/287 (25%), Positives = 119/287 (41%), Gaps = 38/287 (13%) ref|NP_521275.1| PROBABLE SIGNAL PEPTIDE PROTEIN [Ralstonia solanacearum] emb|CAD16942.1| PROBABLE SIGNAL PEPTIDE PROTEIN [Ralstonia solanacearum] Length = 278

1224.2 Best-BlastP=> >nrprot 88% Identities = 212/266 (79%), Positives = 237/266 (89%) emb|CAC35728.1| OXA-29 [Fluoribacter gormanii] Length = 266

1225.2 Best-BlastP=> >nrprot 40% Identities = 90/292 (30%), Positives = 143/292 (48%), Gaps = 26/292 (8%) ref|NP_855062.1| CONSERVED HYPOTHETICAL PROTEIN [Mycobacterium bovis subsp. bovis AF2122/97] emb|CAD94271.1| CONSERVED HYPOTHETICAL PROTEIN [Mycobacterium bovis subsp. bovis AF2122/97] Length = 439

1226.2 Best-BlastP=> >nrprot 55% Identities = 129/350 (36%), Positives = 190/350 (54%), Gaps = 22/350 (6%) ref|NP_629326.1| putative sulfurylase [Streptomyces coelicolor A3(2)] emb|CAC01308.1| putative sulfurylase [Streptomyces coelicolor A3(2)] Length = 392

1227.1 Best-BlastP=> >nrprot 48% Identities = 66/202 (32%), Positives = 109/202 (53%), Gaps = 10/202 (4%) ref|ZP_00107215.1| COG0637: Predicted phosphatase/phosphohexomutase [Nostoc punctiforme] Length = 242

1229.3 Best-BlastP=> >nrprot No Hits found

123.2 Best-BlastP=> >nrprot 59% Identities = 201/462 (43%), Positives = 284/462 (61%), Gaps = 28/462 (6%) ref|NP_253634.1| N-acetylmuramoyl-L-alanine amidase [Pseudomonas aeruginosa PA01] pir||G83028 N-acetylmuramoyl-L-alanine amidase PA4947 [imported] - Pseudomonas aeruginosa (strain PAO1) gb|AAG08332.1|AE004907_10 N-acetylmuramoyl-L-alanine amidase [Pseudomonas aeruginosa PAO1] Length = 475

1230.2 Best-BlastP=> >nrprot 81% Identities = 364/516 (70%), Positives = 429/516 (83%) ref|NP_819831.1| peptide chain release factor 3 [Coxiella burnetii RSA 493] gb|AAO90345.1| peptide chain release factor 3 [Coxiella burnetii RSA 493] Length = 525

1231.3 Best-BlastP=> >nrprot No Hits found

1235.3

Best-BlastP=> >nrprot 28% Identities = 44/137 (32%), Positives = 70/137 (51%), Gaps = 1/137 (0%) ref|NP_800044.1| putative acetyltransferase [Vibrio parahaemolyticus RIMD 2210633] dbj|BAC61877.1| putative acetyltransferase [Vibrio parahaemolyticus] Length = 140

1236.2 Best-BlastP=> >nrprot 65% Identities = 172/176 (97%), Positives = 174/176 (98%) gb|AAM00633.1| unknown [Legionella pneumophila] Length = 176

1237.2 Best-BlastP=> >nrprot 50% Identities = 47/132 (35%), Positives = 73/132 (55%), Gaps = 4/132 (3%) ref|NP_832446.1| Acetyltransferase [Bacillus cereus ATCC 14579] gb|AAP09647.1| Acetyltransferase [Bacillus cereus ATCC 14579] Length = 141

1238.2 Best-BlastP=> >nrprot No Hits found

1239.2 Best-BlastP=> >nrprot No Hits found

124.1 Best-BlastP=> >nrprot 63% Identities = 66/144 (45%), Positives = 95/144 (65%) ref|NP_716232.1| conserved hypothetical protein TIGR00150 [Shewanella oneidensis MR-1] gb|AAN53677.1|AE015507_3 conserved hypothetical protein TIGR00150 [Shewanella oneidensis MR-1] Length = 152

1242.2

Best-BlastP=> >nrprot 82% Identities = 138/191 (72%), Positives = 160/191 (83%) ref|NP_353171.1| AGR_C_216p [Agrobacterium tumefaciens] ref|NP_530843.1| uracil phosphoribosyltransferase [Agrobacterium tumefaciens str. C58 (U. Washington)] sp|Q8UJ06|UPP_AGRT5 Uracil phosphoribosyltransferase (UMP pyrophosphorylase) (UPRTase) pir||C97375 uracil phosphoribosyltransferase (UMP pyrophosphorylase) (uprtase) [imported] - Agrobacterium tumefaciens (strain C58, Cereon) pir||AI2592 uracil phosphoribosyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont) gb|AAK85956.1| AGR_C_216p [Agrobacterium tumefaciens str. C58 (Cereon)] gb|AAL41159.1| uracil phosphoribosyltransferase [Agrobacterium tumefaciens str. C58 (U. Washington)] Length = 209

1244.2 Best-BlastP=> >nrprot No Hits found

1245.1 Best-BlastP=> >nrprot No Hits found

1246.2 Best-BlastP=> >nrprot 22% Identities = 82/298 (27%), Positives = 130/298 (43%), Gaps = 56/298 (18%) pir||T09051 PepA protein - Pseudomonas aeruginosa gb|AAC16023.1| ExoU [Pseudomonas aeruginosa] gb|AAC38269.1| PepA [Pseudomonas aeruginosa] gb|AAP82959.1| type III effector protein [Pseudomonas aeruginosa] Length = 687

1249.2 Best-BlastP=> >nrprot 30% Identities = 37/119 (31%), Positives = 57/119 (47%), Gaps = 9/119 (7%) emb|CAD21525.1| hypothetical protein [Taenia solium] Length = 155

125.2

Best-BlastP=> >nrprot 59% Identities = 227/486 (46%), Positives = 293/486 (60%), Gaps = 4/486 (0%) ref|NP_820087.1| conserved hypothetical protein [Coxiella burnetii RSA 493] gb|AAO90601.1| conserved hypothetical protein [Coxiella burnetii RSA 493] Length = 488

1250.1 Best-BlastP=> >nrprot 64% Identities = 74/160 (46%), Positives = 105/160 (65%), Gaps = 1/160 (0%) ref|NP_820968.1| dihydrofolate reductase [Coxiella burnetii RSA 493] gb|AAO91482.1| dihydrofolate reductase [Coxiella burnetii RSA 493] Length = 161

1251.2 Best-BlastP=> >nrprot 68% Identities = 176/321 (54%), Positives = 222/321 (69%), Gaps = 4/321 (1%) ref|NP_249284.1| pyridoxal phosphate biosynthetic protein PdxA [Pseudomonas aeruginosa PA01] sp|Q9I5U4|PXA1_PSEAE 4-hydroxythreonine-4-phosphate dehydrogenase 1 (4-(phosphohydroxy)-L-threonine dehydrogenase 1) pir||A83572 pyridoxal phosphate biosynthetic protein PdxA PA0593 [imported] - Pseudomonas aeruginosa (strain PAO1) gb|AAG03982.1|AE004495_6 pyridoxal phosphate biosynthetic protein PdxA [Pseudomonas aeruginosa PAO1] Length = 328

1257.5

Best-BlastP=> >nrprot 69% Identities = 242/415 (58%), Positives = 300/415 (72%), Gaps = 5/415 (1%) ref|NP_820851.1| conserved hypothetical protein [Coxiella burnetii RSA 493] gb|AAO91365.1| conserved hypothetical protein [Coxiella burnetii RSA 493] Length = 435

1258.1 Best-BlastP=> >nrprot 67% Identities = 175/312 (56%), Positives = 217/312 (69%), Gaps = 2/312 (0%) ref|ZP_00125152.1| COG0189: Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase) [Pseudomonas syringae pv. syringae B728a] Length = 319

126.2 Best-BlastP=> >nrprot 99% Identities = 470/475 (98%), Positives = 475/475 (100%) emb|CAD42896.1| flagellin [Legionella pneumophila serogroup 1] Length = 475

1260.3 Best-BlastP=> >nrprot 48% Identities = 92/211 (43%), Positives = 126/211 (59%), Gaps = 3/211 (1%) ref|NP_521080.1| PROBABLE LIPOPROTEIN PRECURSOR (VACJ) TRANSMEMBRANE [Ralstonia solanacearum] emb|CAD16666.1| PROBABLE LIPOPROTEIN PRECURSOR (VACJ) TRANSMEMBRANE [Ralstonia solanacearum] Length = 269

1261.2 Best-BlastP=> >nrprot 66% Identities = 286/564 (50%), Positives = 370/564 (65%), Gaps = 8/564 (1%) gb|AAB16855.1| pyruvate decarboxylase [Arabidopsis thaliana] Length = 607

1264.3 Best-BlastP=> >nrprot 39% Identities = 97/291 (33%), Positives = 159/291 (54%), Gaps = 16/291 (5%) ref|NP_899726.1| probable aminopeptidase [Chromobacterium violaceum ATCC 12472] gb|AAQ57736.1| probable aminopeptidase [Chromobacterium violaceum ATCC 12472] Length = 415

1265.5 Best-BlastP=> >nrprot 58% Identities = 119/262 (45%), Positives = 162/262 (61%), Gaps = 6/262 (2%) ref|ZP_00051284.1| COG0656: Aldo/keto reductases, related to diketogulonate reductase [Magnetospirillum magnetotacticum] Length = 291

1267.3 Best-BlastP=> >nrprot 20% Identities = 38/108 (35%), Positives = 55/108 (50%), Gaps = 8/108 (7%) ref|NP_616727.1| conserved hypothetical protein [Methanosarcina acetivorans str. C2A] gb|AAM05207.1| conserved hypothetical protein [Methanosarcina acetivorans str. C2A] Length = 266

1268.2

Best-BlastP=> >nrprot 46% Identities = 82/293 (27%), Positives = 144/293 (49%), Gaps = 33/293 (11%) ref|XP_313252.1| ENSANGP00000010487 [Anopheles gambiae] gb|EAA08759.1| ENSANGP00000010487 [Anopheles gambiae str. PEST] Length = 310

127.4 Best-BlastP=> >nrprot 37% Identities = 32/94 (34%), Positives = 48/94 (51%), Gaps = 6/94 (6%) sp|Q02910|CPN_DROME Calphotin pir||A47282 calcium-binding protein calphotin - fruit fly (Drosophila melanogaster) gb|AAA28405.1| calcium-binding protein Length = 865

1270.2 Best-BlastP=> >nrprot 70% Identities = 73/144 (50%), Positives = 102/144 (70%), Gaps = 2/144 (1%) ref|NP_241923.1| BH1057~unknown conserved protein [Bacillus halodurans] sp|Q9RC41|YA57_BACHD Hypothetical protein BH1057 pir||A83782 hypothetical protein BH1057 [imported] - Bacillus halodurans (strain C-125) dbj|BAA83958.1| YHDE [Bacillus halodurans] dbj|BAB04776.1| BH1057~unknown conserved protein [Bacillus halodurans] Length = 146

1271.3 Best-BlastP=> >nrprot 61% Identities = 35/70 (50%), Positives = 53/70 (75%) sp|P17724|GLB_TETPY Myoglobin (Hemoglobin) pir||A36270 hemoglobin - Tetrahymena pyriformis dbj|BAA03015.1| hemoglobin [Tetrahymena pyriformis] Length = 121

1272.4 Best-BlastP=> >nrprot 46% Identities = 89/318 (27%), Positives = 147/318 (46%), Gaps = 13/318 (4%) ref|NP_253021.1| probable ferredoxin reductase [Pseudomonas aeruginosa PA01] pir||G83104 probable ferredoxin reductase PA4331 [imported] - Pseudomonas aeruginosa (strain PAO1) gb|AAG07719.1|AE004849_6 probable ferredoxin reductase [Pseudomonas aeruginosa PAO1] Length = 308

1273.4 Best-BlastP=> >nrprot 40% Identities = 54/200 (27%), Positives = 95/200 (47%), Gaps = 8/200 (4%) gb|AAM88782.1| hypothetical protein [Photorhabdus luminescens] Length = 247

1275.2 Best-BlastP=> >nrprot 57% Identities = 107/291 (36%), Positives = 171/291 (58%), Gaps = 1/291 (0%) gb|AAM88781.1| MhpE-like protein [Photorhabdus luminescens] Length = 312

1277.2 Best-BlastP=> >nrprot 48% Identities = 100/294 (34%), Positives = 162/294 (55%), Gaps = 11/294 (3%) ref|NP_624285.1| UDP-N-acetylmuramyl tripeptide synthase [Thermoanaerobacter tengcongensis] gb|AAM25889.1| UDP-N-acetylmuramyl tripeptide synthase [Thermoanaerobacter tengcongensis] Length = 879

1278.2 Best-BlastP=> >nrprot 43% Identities = 104/386 (26%), Positives = 176/386 (45%), Gaps = 49/386 (12%) gb|AAN83921.1| hypothetical protein [Aplysia californica] Length = 427

1279.3 Best-BlastP=> >nrprot 76% Identities = 107/173 (61%), Positives = 136/173 (78%) ref|NP_439712.1| hypothetical protein [Haemophilus influenzae Rd] sp|P44255|YFCM_HAEIN Hypothetical protein HI1563 pir||D64036 hypothetical protein HI1563 - Haemophilus influenzae (strain Rd KW20) gb|AAC23212.1| conserved hypothetical protein [Haemophilus influenzae Rd] Length = 178

1280.3 Best-BlastP=> >nrprot 15% Identities = 110/113 (97%), Positives = 110/113 (97%) gb|AAO61471.1| LidA [Legionella pneumophila] Length = 113

1282.3 Best-BlastP=> >nrprot 44% Identities = 155/493 (31%), Positives = 230/493 (46%), Gaps = 37/493 (7%) gb|AAC35592.1| LphB [Legionella pneumophila] Length = 518

1283.2 Best-BlastP=> >nrprot 59% Identities = 180/429 (41%), Positives = 260/429 (60%), Gaps = 9/429 (2%) ref|NP_716339.1| conserved hypothetical protein [Shewanella oneidensis MR-1] gb|AAN53784.1|AE015516_6 conserved hypothetical protein [Shewanella oneidensis MR-1] Length = 433

1284.2 Best-BlastP=> >nrprot 45% Identities = 25/75 (33%), Positives = 43/75 (57%) ref|NP_796782.1| hypothetical protein VP0403 [Vibrio parahaemolyticus RIMD 2210633] dbj|BAC58666.1| hypothetical protein [Vibrio parahaemolyticus] Length = 111

1286.3 Best-BlastP=> >nrprot 19% Identities = 71/260 (27%), Positives = 123/260 (47%), Gaps = 17/260 (6%) gb|AAK19884.1| putative methoxymalonyl CoA synthase [Polyangium cellulosum] Length = 863

1287.3 Best-BlastP=> >nrprot 49% Identities = 278/933 (29%), Positives = 458/933 (49%), Gaps = 56/933 (6%) ref|ZP_00089642.1| hypothetical protein [Azotobacter vinelandii] Length = 973

1288.2 Best-BlastP=> >nrprot 14% Identities = 39/130 (30%), Positives = 67/130 (51%), Gaps = 8/130 (6%) ref|XP_314825.1| ENSANGP00000011098 [Anopheles gambiae] gb|EAA10144.1| ENSANGP00000011098 [Anopheles gambiae str. PEST] Length = 1842

129.3 Best-BlastP=> >nrprot No Hits found

1293.4 Best-BlastP=> >nrprot 62% Identities = 91/182 (50%), Positives = 123/182 (67%), Gaps = 2/182 (1%) ref|ZP_00013245.1| COG2353: Uncharacterized conserved protein [Rhodospirillum rubrum] Length = 201

1294.2 Best-BlastP=> >nrprot 59% Identities = 77/177 (43%), Positives = 105/177 (59%), Gaps = 7/177 (3%) ref|NP_902948.1| probable cytochrome b561 [Chromobacterium violaceum ATCC 12472] gb|AAQ60942.1| probable cytochrome b561 [Chromobacterium violaceum ATCC 12472] Length = 180

1295.1 Best-BlastP=> >nrprot 47% Identities = 64/169 (37%), Positives = 91/169 (53%), Gaps = 1/169 (0%) ref|NP_422185.1| conserved hypothetical protein [Caulobacter crescentus CB15] pir||E87669 conserved hypothetical protein CC3391 [imported] - Caulobacter crescentus gb|AAK25353.1| conserved hypothetical protein [Caulobacter crescentus CB15] Length = 427

1299.3 Best-BlastP=> >nrprot 53% Identities = 150/437 (34%), Positives = 246/437 (56%), Gaps = 6/437 (1%) gb|AAD28727.1|AF112468_6 TraH protein precursor [Salmonella typhimurium] Length = 457

13.1 Best-BlastP=> >nrprot 40% Identities = 57/260 (21%), Positives = 111/260 (42%), Gaps = 28/260 (10%) ref|NP_125735.1| hypothetical protein [Pyrococcus abyssi] pir||G75189 hypothetical protein PAB2321 - Pyrococcus abyssi (strain Orsay) emb|CAB48966.1| Hypothetical protein [Pyrococcus abyssi] Length = 249

130.4 Best-BlastP=> >nrprot 50% Identities = 184/554 (33%), Positives = 274/554 (49%), Gaps = 94/554 (16%) ref|NP_841634.1| Flagellar hook-associated protein 2 [Nitrosomonas europaea ATCC 19718] emb|CAD85506.1| Flagellar hook-associated protein 2 [Nitrosomonas europaea ATCC 19718] Length = 481

1302.2 Best-BlastP=> >nrprot 54% Identities = 129/348 (37%), Positives = 197/348 (56%), Gaps = 10/348 (2%) ref|ZP_00130398.1| COG2200: FOG: EAL domain [Desulfovibrio desulfuricans G20] Length = 367

1303.3 Best-BlastP=> >nrprot 80% Identities = 168/252 (66%), Positives = 207/252 (82%) ref|ZP_00125838.1| COG0024: Methionine aminopeptidase [Pseudomonas syringae pv. syringae B728a] Length = 260

1306.5 Best-BlastP=> >nrprot 98% Identities = 1033/1066 (96%), Positives = 1049/1066 (98%) gb|AAM00612.1| chemiosmotic efflux system protein A-like protein [Legionella pneumophila] Length = 1066

1307.4 Best-BlastP=> >nrprot 97% Identities = 389/418 (93%), Positives = 407/418 (97%) gb|AAM00611.1| proline/glycine betaine transporter-like protein [Legionella pneumophila] Length = 422

131.2 Best-BlastP=> >nrprot 26% Identities = 26/88 (29%), Positives = 42/88 (47%), Gaps = 15/88 (17%) ref|NP_901795.1| hypothetical protein CV2125 [Chromobacterium violaceum ATCC 12472] gb|AAQ59798.1| hypothetical protein CV2125 [Chromobacterium violaceum ATCC 12472] Length = 202

1314.3 Best-BlastP=> >nrprot 57% Identities = 170/474 (35%), Positives = 278/474 (58%), Gaps = 14/474 (2%) ref|ZP_00129665.1| COG1538: Outer membrane protein [Desulfovibrio desulfuricans G20] Length = 494

1317.3 Best-BlastP=> >nrprot No Hits found

1318.4 Best-BlastP=> >nrprot 64% Identities = 110/230 (47%), Positives = 162/230 (70%), Gaps = 10/230 (4%) ref|NP_819966.1| conserved hypothetical protein [Coxiella burnetii RSA 493] gb|AAO90480.1| conserved hypothetical protein [Coxiella burnetii RSA 493] Length = 414

1320.3 Best-BlastP=> >nrprot 67% Identities = 365/741 (49%), Positives = 498/741 (67%), Gaps = 11/741 (1%) ref|ZP_00010561.1| COG0068: Hydrogenase maturation factor [Rhodopseudomonas palustris] Length = 772

1322.2 Best-BlastP=> >nrprot 60% Identities = 173/452 (38%), Positives = 280/452 (61%), Gaps = 6/452 (1%) ref|NP_462786.1| putative MFS family tranport protein (1st mdule) [Salmonella typhimurium LT2] gb|AAL22745.1| putative MFS family tranport protein [Salmonella typhimurium LT2] Length = 475

1323.2 Best-BlastP=> >nrprot No Hits found

1324.4 Best-BlastP=> >nrprot 28% Identities = 76/274 (27%), Positives = 130/274 (47%), Gaps = 35/274 (12%) gb|EAA21537.1| Plasmodium falciparum CDPK2 protein [Plasmodium yoelii yoelii] Length = 565

1325.2 Best-BlastP=> >nrprot 71% Identities = 311/589 (52%), Positives = 419/589 (71%) ref|NP_820983.1| arginyl-tRNA synthetase [Coxiella burnetii RSA 493] gb|AAO91497.1| arginyl-tRNA synthetase [Coxiella burnetii RSA 493] Length = 592

1328.2 Best-BlastP=> >nrprot 55% Identities = 121/344 (35%), Positives = 196/344 (56%), Gaps = 8/344 (2%) ref|NP_819936.1| transporter, putative [Coxiella burnetii RSA 493] gb|AAO90450.1| transporter, putative [Coxiella burnetii RSA 493] Length = 376

1330.5 Best-BlastP=> >nrprot 69% Identities = 372/682 (54%), Positives = 478/682 (70%), Gaps = 4/682 (0%) ref|ZP_00092220.1| COG1200: RecG-like helicase [Azotobacter vinelandii] Length = 1006

1331.3 Best-BlastP=> >nrprot 55% Identities = 125/409 (30%), Positives = 223/409 (54%), Gaps = 18/409 (4%) ref|NP_391052.1| alternate gene name: comB, yufA [Bacillus subtilis] sp|P14203|YUXH_BACSU Hypothetical protein yuxH pir||BVBSCB competence protein ComB (yuxH) - Bacillus subtilis gb|AAA22318.1| B competence protein emb|CAB07900.1| unknown [Bacillus subtilis] emb|CAB15162.1| yuxH [Bacillus subtilis subsp. subtilis str. 168] Length = 409

1332.3 Best-BlastP=> >nrprot 77% Identities = 308/477 (64%), Positives = 377/477 (79%) ref|NP_840694.1| Glycine cleavage system P-protein [Nitrosomonas europaea ATCC 19718] emb|CAD84521.1| Glycine cleavage system P-protein [Nitrosomonas europaea ATCC 19718] Length = 483

1334.3 Best-BlastP=> >nrprot No Hits found

1335.2 Best-BlastP=> >nrprot No Hits found

1336.3 Best-BlastP=> >nrprot 99% Identities = 365/371 (98%), Positives = 369/371 (99%) gb|AAM00605.1| florfenicol efflux pump-like protein [Legionella pneumophila] Length = 371

1338.3 Best-BlastP=> >nrprot 76% Identities = 406/673 (60%), Positives = 510/673 (75%), Gaps = 8/673 (1%) ref|NP_837722.1| methionine tRNA synthetase [Shigella flexneri 2a str. 2457T] gb|AAP17531.1| methionine tRNA synthetase [Shigella flexneri 2a str. 2457T] Length = 677

1339.2 Best-BlastP=> >nrprot 79% Identities = 109/185 (58%), Positives = 151/185 (81%) ref|NP_770170.1| blr3530 [Bradyrhizobium japonicum] dbj|BAC48795.1| blr3530 [Bradyrhizobium japonicum USDA 110] Length = 206

1341.2 Best-BlastP=> >nrprot No Hits found

1342.3 Best-BlastP=> >nrprot 49% Identities = 73/251 (29%), Positives = 130/251 (51%), Gaps = 16/251 (6%) gb|AAM90720.1| TraF [Salmonella typhi] Length = 259

1344.3 Best-BlastP=> >nrprot 14% Identities = 30/113 (26%), Positives = 51/113 (45%), Gaps = 15/113 (13%) ref|ZP_00101173.1| hypothetical protein [Desulfitobacterium hafniense] Length = 367

1345.3 Best-BlastP=> >nrprot 65% Identities = 173/366 (47%), Positives = 239/366 (65%), Gaps = 7/366 (1%) ref|NP_778380.1| conserved hypothetical protein [Xylella fastidiosa Temecula1] gb|AAO28029.1| conserved hypothetical protein [Xylella fastidiosa Temecula1] Length = 381

1346.3 Best-BlastP=> >nrprot 69% Identities = 135/258 (52%), Positives = 183/258 (70%), Gaps = 5/258 (1%) ref|NP_841557.1| Uncharacterized protein family UPF0006 [Nitrosomonas europaea ATCC 19718] emb|CAD85427.1| Uncharacterized protein family UPF0006 [Nitrosomonas europaea ATCC 19718] Length = 254

1353.3 Best-BlastP=> >nrprot 43% Identities = 82/333 (24%), Positives = 152/333 (45%), Gaps = 43/333 (12%) ref|NP_819818.1| multidrug resistance protein [Coxiella burnetii RSA 493] gb|AAO90332.1| multidrug resistance protein [Coxiella burnetii RSA 493] Length = 331

1354.2 Best-BlastP=> >nrprot No Hits found

1356.2 Best-BlastP=> >nrprot 44% Identities = 72/242 (29%), Positives = 117/242 (48%), Gaps = 8/242 (3%) ref|NP_814508.1| amino acid ABC transporter, amino acid-binding/permease protein [Enterococcus faecalis V583] gb|AAO80578.1| amino acid ABC transporter, amino acid-binding/permease protein [Enterococcus faecalis V583] Length = 722

1357.2 Best-BlastP=> >nrprot 71% Identities = 90/146 (61%), Positives = 107/146 (73%) ref|NP_718466.1| YaiI/YqxD family protein [Shewanella oneidensis MR-1] gb|AAN55910.1|AE015727_10 YaiI/YqxD family protein [Shewanella oneidensis MR-1] Length = 151

1359.2 Best-BlastP=> >nrprot 99% Identities = 357/357 (100%), Positives = 357/357 (100%) emb|CAD43479.1| O-acetyltransferase [Legionella pneumophila] Length = 357

1360.6 Best-BlastP=> >nrprot 39% Identities = 112/443 (25%), Positives = 195/443 (44%), Gaps = 75/443 (16%) ref|NP_845547.1| conserved hypothetical protein [Bacillus anthracis str. Ames] gb|AAP27033.1| conserved hypothetical protein [Bacillus anthracis str. Ames] Length = 445

1361.3 Best-BlastP=> >nrprot 46% Identities = 99/312 (31%), Positives = 162/312 (51%), Gaps = 25/312 (8%) ref|NP_250993.1| hypothetical protein [Pseudomonas aeruginosa PA01] pir||A83358 hypothetical protein PA2303 [imported] - Pseudomonas aeruginosa (strain PAO1) gb|AAG05691.1|AE004656_3 hypothetical protein PA2303 [Pseudomonas aeruginosa PAO1] Length = 339

1362.3 Best-BlastP=> >nrprot 96% Identities = 570/607 (93%), Positives = 588/607 (96%) gb|AAK00285.1|AF288536_7 unknown [Legionella longbeachae] Length = 607

1364.2 Best-BlastP=> >nrprot 96% Identities = 505/534 (94%), Positives = 516/534 (96%) gb|AAK00286.1|AF288536_8 possible sensor kinase protein [Legionella longbeachae] Length = 534

1366.3 Best-BlastP=> >nrprot No Hits found

1367.3 Best-BlastP=> >nrprot 58% Identities = 55/123 (44%), Positives = 83/123 (67%) ref|ZP_00047813.1| COG2391: Predicted transporter component [Magnetospirillum magnetotacticum] Length = 155

1368.2 Best-BlastP=> >nrprot No Hits found

1369.2 Best-BlastP=> >nrprot 63% Identities = 58/133 (43%), Positives = 92/133 (69%) ref|NP_800456.1| conserved hypothetical protein [Vibrio parahaemolyticus RIMD 2210633] dbj|BAC62289.1| conserved hypothetical protein [Vibrio parahaemolyticus] Length = 139

1370.5 Best-BlastP=> >nrprot 56% Identities = 66/135 (48%), Positives = 87/135 (64%), Gaps = 2/135 (1%) ref|NP_819361.1| conserved domain protein [Coxiella burnetii RSA 493] gb|AAO89875.1| conserved domain protein [Coxiella burnetii RSA 493] Length = 214

1372.2 Best-BlastP=> >nrprot No Hits found

1373.1 Best-BlastP=> >nrprot No Hits found

1375.2 Best-BlastP=> >nrprot 29% Identities = 51/102 (50%), Positives = 69/102 (67%) ref|NP_790719.1| conserved hypothetical protein [Pseudomonas syringae pv. tomato str. DC3000] gb|AAO54414.1| conserved hypothetical protein [Pseudomonas syringae pv. tomato str. DC3000] Length = 249

1376.2 Best-BlastP=> >nrprot 40% Identities = 38/97 (39%), Positives = 53/97 (54%), Gaps = 3/97 (3%) ref|ZP_00030779.1| hypothetical protein [Burkholderia fungorum] Length = 182

1377.5 Best-BlastP=> >nrprot No Hits found

1378.5 Best-BlastP=> >nrprot No Hits found

1380.3 Best-BlastP=> >nrprot 15% Identities = 51/177 (28%), Positives = 81/177 (45%), Gaps = 26/177 (14%) ref|ZP_00071821.1| COG3240: Phospholipase/lecithinase/hemolysin [Trichodesmium erythraeum IMS101] Length = 418

1385.2 Best-BlastP=> >nrprot 40% Identities = 61/216 (28%), Positives = 98/216 (45%), Gaps = 23/216 (10%) gb|AAK00282.1|AF288536_4 unknown [Legionella longbeachae] Length = 302

1386.3 Best-BlastP=> >nrprot 57% Identities = 56/108 (51%), Positives = 72/108 (66%) gb|AAL17787.1|AF361470_8 hypothetical protein [Rhizobium leguminosarum bv. trifolii] Length = 115

1387.2 Best-BlastP=> >nrprot 58% Identities = 191/434 (44%), Positives = 267/434 (61%), Gaps = 11/434 (2%) ref|ZP_00030193.1| COG1538: Outer membrane protein [Burkholderia fungorum] Length = 509

1388.2 Best-BlastP=> >nrprot No Hits found

139.5 Best-BlastP=> >nrprot 10% Identities = 77/305 (25%), Positives = 150/305 (49%), Gaps = 30/305 (9%) gb|AAK01145.2| 200 kDa immunoreactive glycoprotein [Ehrlichia canis] Length = 1421

1390.5 Best-BlastP=> >nrprot No Hits found

1391.1

Best-BlastP=> >nrprot 73% Identities = 119/193 (61%), Positives = 151/193 (78%), Gaps = 1/193 (0%) ref|NP_820272.1| proton transporter, putative [Coxiella burnetii RSA 493] gb|AAO90786.1| proton transporter, putative [Coxiella burnetii RSA 493] Length = 200

1392.2 Best-BlastP=> >nrprot 41% Identities = 31/81 (38%), Positives = 45/81 (55%), Gaps = 2/81 (2%) ref|ZP_00092569.1| COG3293: Transposase and inactivated derivatives [Azotobacter vinelandii] Length = 249

1394.2 Best-BlastP=> >nrprot 63% Identities = 55/150 (36%), Positives = 95/150 (63%), Gaps = 1/150 (0%) ref|ZP_00031963.1| COG0517: FOG: CBS domain [Burkholderia fungorum] Length = 150

1396.3 Best-BlastP=> >nrprot 75% Identities = 204/328 (62%), Positives = 250/328 (76%) ref|NP_484922.1| alcohol dehydrogenase [Nostoc sp. PCC 7120] pir||AD1916 alcohol dehydrogenase [imported] - Nostoc sp. (strain PCC 7120) dbj|BAB72836.1| alcohol dehydrogenase [Nostoc sp. PCC 7120] Length = 328

1397.2 Best-BlastP=> >nrprot 52% Identities = 56/149 (37%), Positives = 86/149 (57%), Gaps = 4/149 (2%) ref|NP_771277.1| blr4637 [Bradyrhizobium japonicum] dbj|BAC49902.1| blr4637 [Bradyrhizobium japonicum USDA 110] Length = 173

1398.2

Best-BlastP=> >nrprot 56% Identities = 261/666 (39%), Positives = 375/666 (56%), Gaps = 31/666 (4%) ref|NP_901612.1| probable peptidase [Chromobacterium violaceum ATCC 12472] gb|AAQ59616.1| probable peptidase [Chromobacterium violaceum ATCC 12472] Length = 660

1399.3 Best-BlastP=> >nrprot 47% Identities = 135/479 (28%), Positives = 236/479 (49%), Gaps = 16/479 (3%) ref|NP_460415.1| putative POT family, peptide transport protein [Salmonella typhimurium LT2] gb|AAL20374.1| putative POT family peptide transport protein [Salmonella typhimurium LT2] Length = 501

1400.2 Best-BlastP=> >nrprot 71% Identities = 245/484 (50%), Positives = 346/484 (71%), Gaps = 8/484 (1%) ref|NP_820418.1| NADH dehydrogenase I, N subunit [Coxiella burnetii RSA 493] gb|AAO90932.1| NADH dehydrogenase I, N subunit [Coxiella burnetii RSA 493] Length = 482

1402.2 Best-BlastP=> >nrprot 73% Identities = 289/499 (57%), Positives = 367/499 (73%), Gaps = 1/499 (0%) ref|NP_820419.1| NADH dehydrogenase I, M subunit [Coxiella burnetii RSA 493] gb|AAO90933.1| NADH dehydrogenase I, M subunit [Coxiella burnetii RSA 493] Length = 506

1405.2 Best-BlastP=> >nrprot 98% Identities = 292/296 (98%), Positives = 293/296 (98%) emb|CAB43070.1| DjlA protein [Legionella pneumophila] Length = 296

1406.2 Best-BlastP=> >nrprot 97% Identities = 401/419 (95%), Positives = 408/419 (97%) emb|CAB43071.1| 3-Deoxy-D-manno-oct-2-ulosonic acid transferase [Legionella pneumophila] Length = 419

1407.2 Best-BlastP=> >nrprot No Hits found

1409.3 Best-BlastP=> >nrprot 60% Identities = 102/245 (41%), Positives = 157/245 (64%), Gaps = 3/245 (1%) ref|NP_819783.1| competence lipoprotein ComL, putative [Coxiella burnetii RSA 493] gb|AAO90297.1| competence lipoprotein ComL, putative [Coxiella burnetii RSA 493] Length = 255

141.4 Best-BlastP=> >nrprot 25% Identities = 116/527 (22%), Positives = 221/527 (41%), Gaps = 72/527 (13%) pir||T18296 myosin heavy chain - Entamoeba histolytica gb|AAB48065.1| myosin heavy chain [Entamoeba histolytica] Length = 2139

1410.2 Best-BlastP=> >nrprot 50% Identities = 94/303 (31%), Positives = 159/303 (52%), Gaps = 8/303 (2%) ref|NP_931129.1| recombinaison associated protein [Photorhabdus luminescens subsp. laumondii TTO1] emb|CAE16301.1| recombinaison associated protein [Photorhabdus luminescens subsp. laumondii TTO1] Length = 303

1412.5 Best-BlastP=> >nrprot 61% Identities = 251/573 (43%), Positives = 359/573 (62%), Gaps = 17/573 (2%) ref|ZP_00110863.1| hypothetical protein [Nostoc punctiforme] Length = 578

1413.2 Best-BlastP=> >nrprot 36% Identities = 99/349 (28%), Positives = 168/349 (48%), Gaps = 18/349 (5%) ref|NP_284204.1| putative acyl-CoA ligase [Neisseria meningitidis Z2491] pir||D81839 probable acyl-CoA ligase (EC 6.2.1.-) NMA1482 [imported] - Neisseria meningitidis (strain Z2491 serogroup A) emb|CAB84715.1| putative acyl-CoA ligase [Neisseria meningitidis Z2491] Length = 517

1414.1 Best-BlastP=> >nrprot 52% Identities = 24/70 (34%), Positives = 41/70 (58%), Gaps = 2/70 (2%) ref|ZP_00036213.1| hypothetical protein [Enterococcus faecium] Length = 77

1415.4 Best-BlastP=> >nrprot 48% Identities = 75/253 (29%), Positives = 121/253 (47%), Gaps = 30/253 (11%) ref|ZP_00012801.1| hypothetical protein [Rhodopseudomonas palustris] Length = 255

1417.1 Best-BlastP=> >nrprot 66% Identities = 64/129 (49%), Positives = 88/129 (68%) ref|ZP_00066647.1| hypothetical protein [Microbulbifer degradans 2-40] Length = 164

1418.3 Best-BlastP=> >nrprot 71% Identities = 295/592 (49%), Positives = 421/592 (71%), Gaps = 2/592 (0%) ref|NP_820066.1| transporter, putative [Coxiella burnetii RSA 493] gb|AAO90580.1| transporter, putative [Coxiella burnetii RSA 493] Length = 593